Journal of MOLECULAR BIOLOGY

EDITOR-IN-CHIEF: J. C. KENDREW

Volume III, Number 3, June 1961

ACADEMIC PRESS LTD., LONDON
ACADEMIC PRESS INC., NEW YORK

JOURNAL OF MOLECULAR BIOLOGY

EDITOR-IN-CHIEF

J. C. KENDREW, Peterhouse, Cambridge, England

ASSISTANT EDITOR

S. BRENNER, M.R.C. Unit, Cavendish Laboratory, University of Cambridge, England.

EDITORIAL BOARD

P. DOTY, Department of Chemistry, Harvard University, Cambridge, Mass., U.S.A.
A. F. HUXLEY, Department of Physiology, University College, London, England.
M. MESELSON, Biological Laboratories, Harvard University, Cambridge, Mass., U.S.A.
R. L. SINSHEIMER, Division of Biology, California Institute of Technology, Pasadena, Cal., U.S.A.
J. D. WATSON, Biological Laboratories, Harvard University, Cambridge, Mass., U.S.A.
M. H. F. WILKINS, Physics Department, King's College, London, England.

ADVISORY BOARD

S. BENZER, Purdue University, U.S.A.
M. CALVIN, University of California, Berkeley, U.S.A.
F. H. C. CRICK, University of Cambridge, England.
J. F. DANIELLI, King's College, London, England.
A. ENGSTRÖM, Karolinska Institutet, Stockholm, Sweden.
A. D. HERSHEY, Carnegie Institute, Cold Spring Harbor, U.S.A.
F. JACOB, Pasteur Institute, Paris, France.
A. KORNBERG, Stanford University, Stanford, California, U.S.A.
C. LEVINTHAL, Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.
S. E. LURIA, Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.
D. MAZIA, University of California, Berkeley, U.S.A.
P. B. MEDAWAR, University College, London, England.
L. E. ORGEL, University of Cambridge, England.
M. F. PERUTZ, University of Cambridge, England.
R. D. PRESTON, University of Leeds, England.
G. N. RAMACHANDRAN, University of Madras, India.
A. RICH, Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.
G. SCHRAMM, Max-Planck Institut, Tübingen, Germany.
F. S. SJØSTRAND, Karolinska Institutet, Stockholm, Sweden.
I. WATANABE, Institute for Virus Research, Kyoto University, Japan.
R. C. WILLIAMS, University of California, Berkeley, U.S.A.

Published by Academic Press Inc. (London) Limited

Volume III, 1961: 128*s.* 6*d.* ($18.00)

Subscription orders should be sent to Academic Press Inc. (London) Limited, 17 Old Queen Street, London, S.W.1, with the exception of those originating in the U.S.A., Canada, Central America, and South America; these should be sent to Academic Press Inc., 111 Fifth Avenue, New York 3, N.Y.

Copyright © 1961, by Academic Press Inc. (London) Limited

J. Mol. Biol. (1961) **3**, 229–240

The Structure of DNA and RNA in the Water-Mercury Interface

I. R. Miller †

Polymer Department, The Weizmann Institute of Science, Rehovot, Israel

(*Received 23 August 1960*)

The interaction of nucleic acids with a charged mercury surface has been investigated. The investigation was carried out by measuring the effect of the nucleic acid on the differential capacity of the electrical double layer between a polarized mercury surface and an 0·1 N-NaCl solution containing various concentrations of nucleic acid. The lowering of the capacity by adsorbed nucleic acid gives an indication of the adherence to the surface of the different molecular residues (sugar, purine, ionized groups) of the nucleic acid.

The kinetics of adsorption were used to determine the surface concentration. This was done after discussing evidence for a diffusion controlled mechanism of adsorption. The lowering of the differential capacity is proportional to the surface concentration for partly covered surfaces and reaches a constant value after the surface becomes fully covered. Additional adsorption by building further layers does not affect the differential capacity. This makes it possible to determine limiting values of areas per nucleotide, when full surface coverage is reached.

The areas determined in this way are: 96 Å^2/nucleotide for *E. coli* RNA, and 93 Å^2/nucleotide for denatured herring sperm DNA, irrespective of the polarization of the mercury surface. At positively charged mercury surfaces, the area per nucleotide for native salmon sperm DNA is of the same order (86 Å^2), but at negatively charged surfaces the area is only 35 Å^2/nucleotide. The results indicate that native DNA preserves its double helical structure when adsorbed on a negatively charged surface, whereas unfolding occurs on a positively charged mercury surface.

An unfolding mechanism of the native DNA in the surface is discussed. According to the proposed model, the unfolding of the helix starts at that point in the chain which reaches the surface first and proceeds by rotation of the two double helical cylinders which protrude into the solution about their axis. The torque for rotating the cylinders is created by the binding of the nucleotides at the unravelling point which pass into an adsorbed state on the surface.

1. Introduction

The physical properties of solutions of nucleic acids have been extensively investigated by different authors, the standard methods such as light scattering (Reichmann, Rice, Thomas & Doty, 1954), viscosity (Pouyet, 1952; Eisenberg, 1957), sedimentation (Hermans & Hermans, 1959) etc. being used. Perhaps the most inspiring of all the published works on nucleic acid solutions are those dealing with the transition from the helical arrangement of the nucleic acids to the unfolded one. In this connection the potentiometric titration of DNA (Gulland, Jordan & Taylor, 1947; Cox & Peacocke, 1956) and of RNA (Cox, Jones, Marsh & Peacocke, 1956; Littauer & Eisenberg, 1959) ought to be mentioned. A more complete study of strand separation and recombination by elevating and slowly lowering the temperature was carried out

† Present address : School of Mines, Columbia University, New York.

by Doty and his collaborators (Doty, Marmur, Eisner & Schildkraut, 1960). Strand separation was followed by the change in the optical absorbance at 260 mμ as well as by density gradient ultracentrifugation and electron microscopy.

The ultimate aim of the study of the transition process is to find some clue to the mechanism of biological processes involving nucleic acids. As various kinds of interfaces including "microfaces" of protein molecules are common in the living cell, it was felt that an investigation of the interaction of nucleic acids with surfaces might be instructive. Although a mercury surface is obviously very far from resembling a biological surface, it may serve as a convenient tool for investigating the interaction of biological polyelectrolytes with charged surfaces.

In the present work the behaviour of nucleic acids on a polarized mercury surface is described. This behaviour was investigated by measuring changes in the electrical properties of the mercury surface produced by the adsorption of nucleic acids. Of all the electrical surface properties, the differential capacity of the electrical double layer is the most conveniently measured. The entry of the solute into the mercury surface affects the capacity by changing the dielectric constant, by interfering with the close approach of the ions to the surface (Miller & Grahame, 1956) and by introducing new charges (Miller & Grahame, 1957). Changes in the differential capacity due to the adsorption of nucleic acids thus give information about the number and kind (purine, sugar, charge) of residues attached to the surface. In addition, the kinetics of adsorption enable us to calculate the surface concentration or the area occupied by a molecular unit in the surface. From the information obtained this way the structure of the molecule in the surface can be deduced. An important advantage of this technique is the ease with which the surface polarization of the mercury can be altered simply by changing the potential difference between it and a reversible electrode placed in the aqueous solution. The changes in the potential are accompanied by a variation in the charge density of the mercury surface (Grahame, 1947).

2. Experimental

(a) *Measurements*

The differential capacity of the electrical double layer at the interface between mercury and an 0·1 N-NaCl solution containing various concentrations of nucleic acid has been measured in an impedance bridge at 25°C. The apparatus is described in detail elsewhere (Grahame, 1949): a brief description follows. Mercury droplets are allowed to form by dripping from the tip of a thin-walled capillary placed within the solution to be investigated and an electrical double layer arises at the interface. The mercury droplet is surrounded by a platinum gauze sphere which serves as a second electrode for the impedance measurements and also as an auxilliary potential-fixing electrode for the polarization circuit. Its potential relative to the solution was monitored with a 1 N-calomel electrode and measurements were reckoned relative to the latter. Balance in the bridge is reached at a certain measured time after the droplet starts growing. As the rate of flow of mercury was constant, the area of the mercury droplet was determined by its age.

(b) *Calculation of the surface concentration*

It is assumed that the adsorption is diffusion-controlled and that any molecule reaching the surface becomes instantaneously adsorbed so that the concentration of the nucleic acids in the layer adjacent to the surface is practically zero in comparison with the bulk concentration C_b. This assumption is justified since, even at extremely low concentration of nucleic acid, saturation values of differential capacity are reached

at reasonably long drop times. If insufficient time is allowed for the adsorption process to run to completion the surface will be only partly covered and the amount adsorbed Γ_t at a drop age t will be given by the following expression (Koryta, 1953; Delahay & Trachtenberg, 1957):

$$\Gamma_t = 0{\cdot}745\, D^{\frac{1}{2}} t^{\frac{1}{2}} C_b \qquad (1)$$

where D is the diffusion coefficient. Saturation surface coverage Γ_{sat} will occur at a certain t_{sat} for each given C_b and at a certain C_{bsat} for each given t.

Any measured surface property is proportional to Γ_t and thus to C_b if measurements are made at constant t, and to $t^{\frac{1}{2}}$ if measurements are made at constant C_b. It was shown that, in the case of synthetic polyelectrolytes, surface tension, surface charge density and differential capacity behave in this way (Miller & Grahame, 1960; Miller, 1960, 1961).

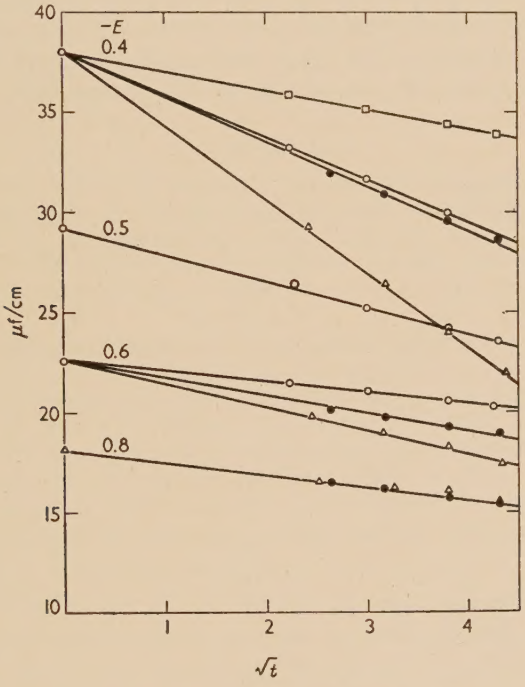


FIG. 1. Differential capacity plotted against the square root of the drop age $t^{\frac{1}{2}}$ at different potential and in the presence of various nucleic acids:

○ — native salmon sperm DNA $C_b = 0{\cdot}08$ mg/ml.;
□ — native salmon sperm DNA $C_b = 0{\cdot}04$ mg/ml.;
△ — denatured herring sperm DNA $C_b = 0{\cdot}034$ mg/ml.;
● — *E. coli* RNA $C_b = 0{\cdot}028$ mg/ml.

Similar relations between differential capacity and $t^{\frac{1}{2}}$ and C_b hold also in the present case with native salmon sperm DNA, denatured herring sperm DNA and *E. coli* RNA (see below). The linear dependence of the capacity on $t^{\frac{1}{2}}$ at constant concentration of RNA and of the two samples of DNA in the presence of 0·1 N-NaCl at pH 6 is shown in Fig. 1. It should be noticed that the straight lines when extrapolated to zero time give the capacity values of pure salt solutions. It is evident therefore that the differential capacity in this region is proportional to surface concentration. The linear

dependence of the capacity on the concentration C_b, up to concentrations at which most of the surface is covered, gives another indication that the adherence to the surface, including the necessary structural rearrangement of the molecules, is very fast and that diffusion to the surface is the rate controlling step.

At high polymer concentration, when the surface becomes fully covered, there is no or only a minute variation in differential capacity with further increase in concentration. This becomes clear when we bear in mind that only the residues which adhere to the mercury surface affect the capacity value significantly (Miller & Grahame, 1956). The number of adhering units remains constant after the surface has become fully covered.

An extrapolation of the differential capacity at low concentrations, where it is proportional to surface coverage, to a concentration at which the surface just becomes fully covered is thus possible. The intercept of the linear plot of differential capacity against concentration for dilute solutions with the almost constant capacity values at high concentrations gives the limiting concentration at which the surface becomes fully covered at the drop time in question. From this intercept and with the aid of equation (1) the surface concentration Γ_{sat} of the fully covered surface could be calculated.

(c) *Materials*

E. coli RNA was obtained from Dr. U. Littauer of the Biochemistry Department. It consisted of two components in about equal amounts, with sedimentation coefficients of 23 s and 16 s. According to Kurland (1960) the corresponding molecular weights are $1{\cdot}1\times10^6$ and $5{\cdot}6\times10^5$. The diffusion coefficients calculated by the Svedberg formula are 11×10^{-8} and 15×10^{-8} respectively. For calculation of surface concentrations the average value was taken.

Native salmon sperm DNA was obtained from Professor P. Doty of Harvard University. Its molecular weight was about 8×10^6 and its diffusion coefficient about $1{\cdot}75\times10^{-8}$.

Herring sperm DNA was bought from the California Foundation for Biochemical Research (allegedly a purified high molecular sample). Variation of optical density at 260 mμ with pH indicated that it was at least partly denatured. The denaturation was completed by dissolving in distilled water and heating for a few minutes at 80°C. The diffusion coefficient was determined from the measured sedimentation coefficient and intrinsic viscosity using the following relation (Sheraga & Mandelkern, 1953):

$$D = \beta^{3/2}(1 - \bar{v}\rho)^{\frac{1}{2}}kT/S^{\frac{1}{2}}N_{Av}^{\frac{1}{2}}\eta_0^{3/2}[\eta]^{\frac{1}{2}} \qquad (2)$$

Here S is the sedimentation coefficient at 25°C, N_{Av} is the Avogadro number, η_0 is the viscosity of water, $[\eta]$ is the intrinsic viscosity, β was assumed to be $2{\cdot}6\times10^6$ and $(1 - \bar{v}\rho) = 0{\cdot}45$.

3. Results

(a) *Differential capacity curves*

A considerable amount of information about the configuration of nucleic acids in the mercury surface can be derived from the plot of the differential capacity against potential in the presence of nucleic acid. The differential capacity curves in the presence of different concentrations of denatured and native DNA are shown in Figs. 2 and 3 respectively. The differential capacity curve in the presence of RNA at full surface coverage is similar to that in the presence of denatured DNA. At concentrations and drop ages where the surface is not yet fully covered the differential capacity curves are a function of the diffusion coefficient of the adsorbing nucleic acid.

In Figs. 2 and 3 an adsorption peak at about $-$ 1·3 V relative to the 1 N-calomel electrode is evident. At more negative potentials the nucleic acids do not interact

with the surface and do not affect the capacity of the double layer. In general, denatured DNA and also RNA reduce the capacity of the electrical double layer to a somewhat larger extent than native DNA. The difference between native and denatured DNA in this respect is larger at potentials more negative than about — 0·6 v relative which is close to the electrocapillary maximum (zero charge on the mercury surface), whereas at more positive potentials the differences are relatively small.

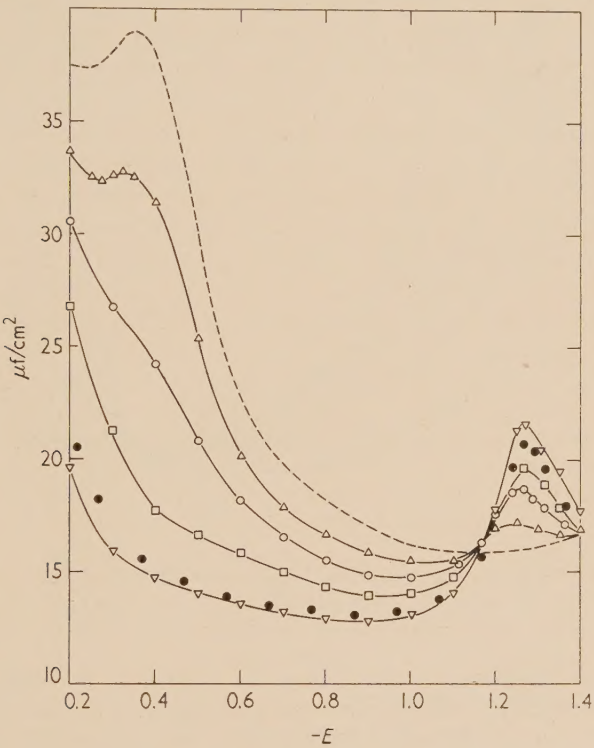


FIG. 2. Differential capacity *versus* potential in the presence of 0·1 N-NaCl and of different concentrations of denatured herring sperm DNA:
- - - - — no DNA; △ — C_b = 0·017 mg/ml.; ○ — C_b = 0·034 mg/ml.;
□ — C_b = 0·068 mg/ml.; ● — C_b = 0·108 mg/ml.

The differences in the surface behaviour of RNA and of native and denatured DNA are even better illustrated in Fig. 4. Here C/C_0, the differential capacity divided by the differential capacity in the presence of the supporting electrolyte only (0·1 N-NaCl), is plotted against potential. The significance of the plot in Fig. 4 will appear from the following consideration. The measured differential capacity C in presence of a surface active material is given by (Melik-Gaikazyan, 1952; Frumkin & Melik-Gaikazyan, 1951):

$$C = \theta C' + (1 - \theta)C_0 + C_{ads} \tag{3}$$

$$C_{ads} = (\partial q/\partial \Gamma)E(\partial \Gamma/\partial E)q$$

where θ is the degree of coverage of the surface by adhering nucleotide units, C' is the differential capacity for these areas where a nucleotide sticks to the surface, C_0 is the differential capacity of the uncovered surface and C_{ads} is the increment in differential capacity due to the variation of the surface charge density q at constant

potential E with change in the surface concentration Γ. C_{ads} will make large contribution only at the adsorption peak.

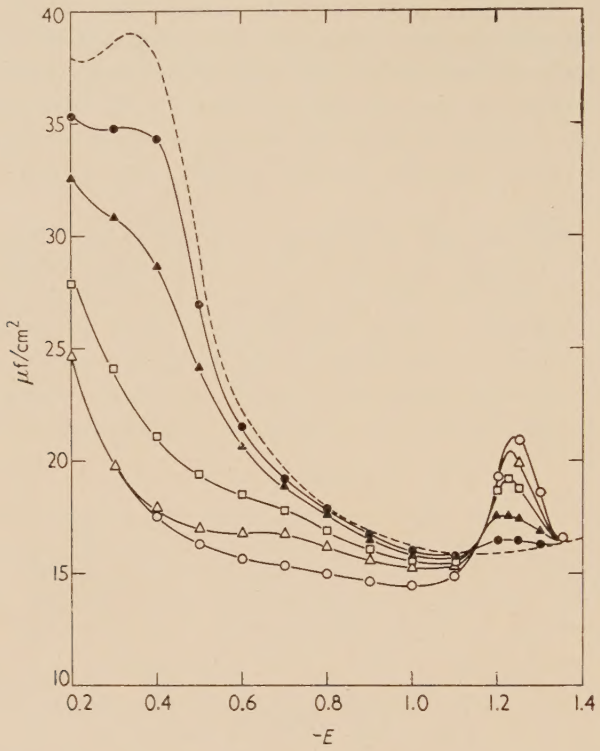


FIG. 3. Differential capacity *versus* potential in the presence of 0·1 N-NaCl and of different concentrations of native salmon sperm DNA:

- - - - — no DNA; ● — C_b = 0·04 mg/ml.; ▲ — C_b = 0·08 mg/ml.;
□ — C_b = 0·16 mg/ml.; △ — C_b = 0·28 mg/ml.; ○ — C_b = 0·60 mg/ml.

Equation (3) can be rewritten

$$C/C_0 = \theta C'/C_0 + 1 - \theta + C_{ads}/C_0 \qquad (4)$$

In Fig. 4 the extrapolated dotted lines eliminate the contribution of C_{ads}/C_0 and represent the "true" capacity (Frumkin & Melik-Gaikazyan, 1951) which depends only on the structure of the double layer and not on the adsorption process. Equation (4) shows that the "true" capacity divided by C_0 is always smaller than 1 as $C' < C_0$. The limiting value for C/C_0 when $C'/C_0 = 0$ is $1 - \theta$.

Assuming that $C_{ads} = 0$ equation (3) can also be written as

$$\Delta C = C_0 - C = \theta(C_0 - C') \qquad (5)$$

It follows from equation (5) that a comparison of ΔC values for different nucleic acids can give a measure for their relative adherence to the surface provided $(C_0 - C')$ can be assumed constant. The effect of the adsorbed residues on the local dielectric constant in the surface should not be confused with the effect of the whole polyelectrolyte molecule on the dielectric constant of the bulk of the water. The latter results probably from the influence of the macromolecule as a whole on the water

structure (Jacobson, 1955). The effect of the adhering residues on the differential capacity is connected with their hydrophobic character (Gorodetskaya & Frumkin, 1938). The more hydrophobic purine and pyrimidine groups would be expected to lower the capacity of the double layer more strongly than the more hydrophilic sugar groups.

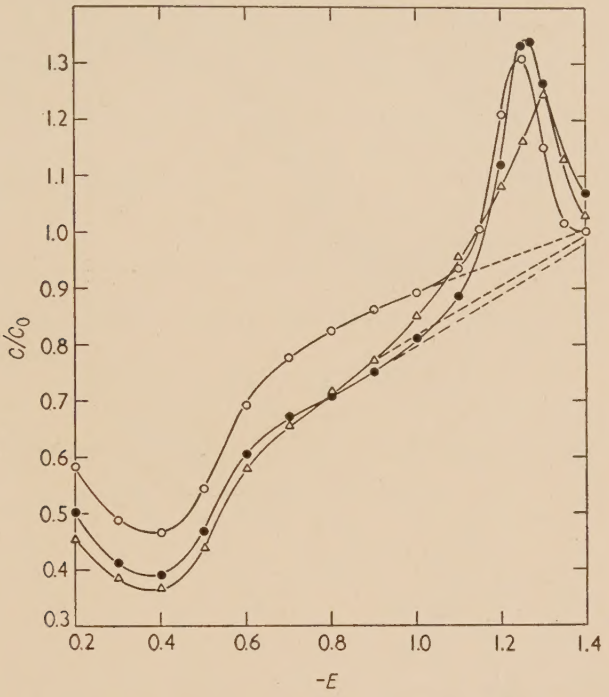


FIG. 4. Reduced differential capacity (C/C_0) *versus* potential at saturation concentrations of nucleic acids:
○ — native salmon sperm DNA; ● — denatured herring sperm DNA; △ — *E. coli* RNA

The adherence of the purine or pyrimidine groups to the surface requires a certain degree of unfolding of the molecule. For geometric reasons on the other hand an increased degree of unfolding increases the possibilities for more complete coverage on the surface. A decrease in C' is therefore bound up with an increase in θ and they both indicate an increase in the degree of unfolding in the surface. On the basis of this reasoning Fig. 5, which gives that ratio in surface coverage of RNA and denatured DNA relative to the native DNA, indicates that native DNA is about as much unfolded at potentials more positive than — 0·6 v as denatured DNA and RNA but that it is much less unfolded at the more negative potentials. This suggests that at the negative potential native DNA preserves its natural double helical structure while it is opened up at the more positive potentials.

(b) *Surface concentration*

Even more convincing evidence for the configuration in the surface of the different nucleic acids can be obtained from the limiting surface concentration. The limiting surface concentrations Γ_{sat} have been derived from the adsorption kinetics by a

method described in the experimental part. Plots of differential capacity against DNA concentrations at constant drop time and at different constant potentials are given

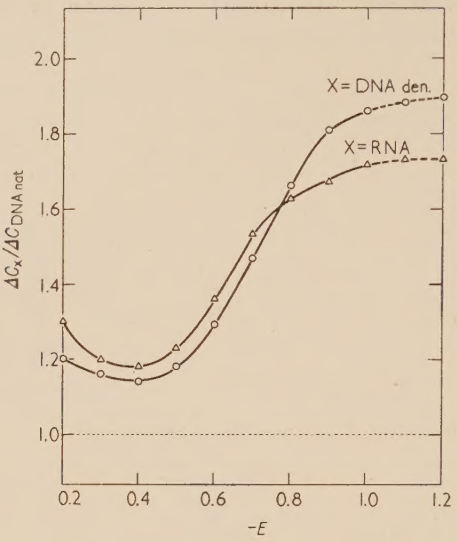


FIG. 5. Relative lowering of the differential capacity by nucleic acids plotted against potential.

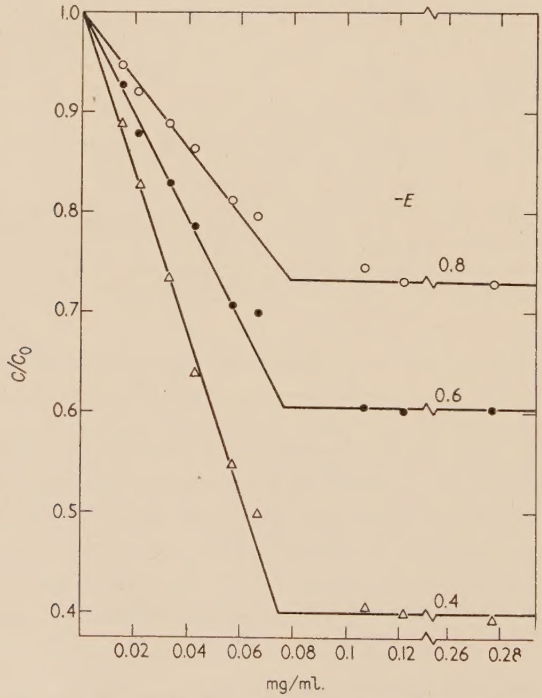


FIG. 6. Reduced differential capacity (C/C_0) plotted at differential potentials against concentration of denatured herring sperm DNA. Drop age = 14·5 sec.

in Figs. 6 and 7 for denatured and native DNA respectively. The intercepts give the critical concentrations at which the mercury surface of the given drop age (14·5 sec) is just fully covered. This critical concentration is independent of potential for RNA

and denatured DNA. Native DNA however shows a sharp transition in the critical concentration value around — 0·6 v. The surface concentration of the fully covered surface is now calculated from this critical concentration with the aid of equation (1).

Table 1

Nucleic acid	Potential relative to 1 N-calomel electrode	Surface conc. (mg/cm²) × 10⁴ Γ_{sat}	Area/nucleotide (Å²)
E. coli RNA	all	0·59	96
Herring sperm RNA (denatured)	all	0·6	93
Salmon sperm DNA (native)	— 0·5 to — 0·1	0·65	86
	— 0·7 to — 1·1	1·6	35

The surface concentration and the areas per nucleotide calculated in this way are summarized in Table 1.

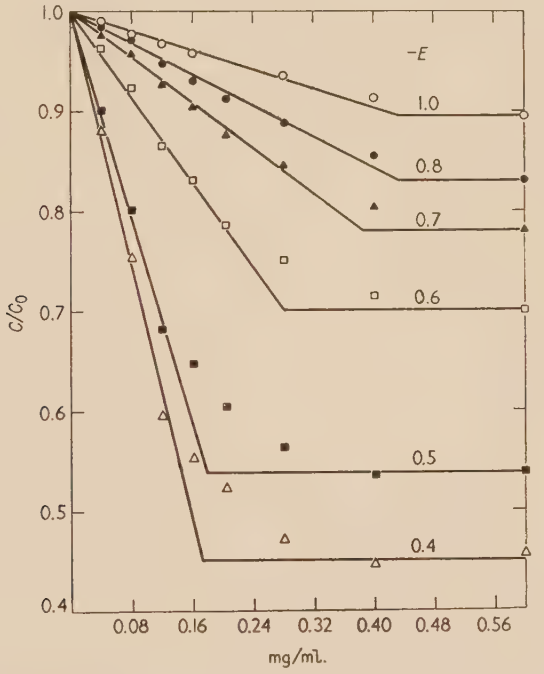


Fig. 7. Reduced differential capacity (C/C_0) plotted at different potentials against the concentration of native salmon sperm DNA. Drop age = 14·5 sec.

As seen from Fig. 7 and Table 1 the limiting areas of the native DNA vary from 35 Å²/nucleotide when the mercury surface is positively charged to 86 Å²/nucleotide when the surface becomes negatively charged. The area of 35 Å²/nucleotide at negative potentials is in good agreement with the area occupied by a nucleotide in a double helix (Watson & Crick, 1953) having a diameter of 20 Å² and a nucleotide pair every 3·4 Å along it. The conclusion is that the double helical form of native DNA is preserved

when adsorbed on a negatively charged surface. It is possible, however, that some of the hydrogen bonds between the purines and pyrimidines are broken. The purine and pyrimidine groups can then turn around the sugar phosphate bonds and adhere to the surface. This possibility is discussed in another publication (Miller, see p. 357).

The much higher surface areas per nucleotide of RNA, of denatured DNA and of native DNA, at positive potentials, indicate that the molecules are unfolded in the surface. As a large number of different configurations of nucleotides could correspond to a limiting surface area of about 90 Å^2/nucleotide it is impossible as yet to draw any conclusions concerning the relation between this area and the surface structure of the nucleic acids. It is likely however that the lower surface area of native DNA at positive potentials (86 Å^2) as compared with the areas of RNA or denatured DNA indicates that the unfolding of the native DNA has not gone to completion.

4. Discussion

The unfolding mechanism

The unfolding of the DNA molecule at the positively charged surface is a very fast process. The experiment showed that the unfolding of a molecule in the surface takes considerably less than 5 sec, although the rate of the process could not be estimated as yet.

It seems clear than an unfolding in the manner proposed by Longuet-Higgins & Zimm (1960) cannot take place in the case of a molecule attached to the surface as the unwinding strands would collide with the surface. An alternative mechanism which could account for the present case is therefore proposed. We assume that the section of the DNA molecule which reaches the surface immediately starts unfolding at the point of contact owing to surface forces. The torque created by the adsorption forces now turns the two double helical cylinders attached to the two strands spreading separately in the surface about their own longitudinal axis. The essential differences between this model and that of Longuet-Higgins & Zimm (1960) are thus the following:

(i) that we assume that the necessary rotation required topologically in order to free the strands involves the rotation relative to the medium not of the free strands but of the still helical parts of the molecule. This will occur at the expense of much less frictional torque particularly at stages where large sections of the chains have already been unwound. In the surface it would be the only possibility.

(ii) that we assume that unwinding can also start from an interior defect in the double helix and that each half of the so divided double chains unravels separately.

The order of magnitude of the unraveling torque may be estimated approximately from the following consideration: the free energy gain per nucleotide in going from the double helix to the dissociated form, with the nucleotides adsorbed into the surface, is given essentially by the free energy of the dissociation plus the free energy of adsorption per nucleotide. The former is about $-3{\cdot}5\times10^{-14}$ erg/nucleotide (Warner & Breslow, 1958) whereas the latter may be estimated from the decrease in surface free energy of the saturated surface (15 erg/cm^2) by multiplying this by the known area per nucleotide (90×10^{-16} cm^2). The free energy per nucleotide in passing, while on the surface, from a position in the double helix to an adjacent adsorption site is thus given by $15\times90\times10^{-16} - 3{\cdot}5\times10^{-14} = 10^{-13}$ erg/nucleotide. We now assume that this change of free energy occurs along a constant slope, i.e. the force acting upon the

nucleotide was constant throughout the process. The change in free energy per nucleotide thus equals the product of this force and the distance through which it has moved. The latter is of the order of the radius r of the double helix so that we may write

$$fr = 10^{-13} \text{ erg/nucleotide}$$

The torque which the moving nucleotide transmits to the still helical and non-adsorbed portions of the molecule is also obviously fr and if we assume that about 4 nucleotide units will at any instant be in the process of transition the whole applied torque τ is given by $\tau = 4fr = 4\times10^{-13}$. This may be put equal to the hydrodynamic resistance of the rotating cylinder and we have

$$\tau = 4\times10^{-13} = 4\pi l\eta_0\omega r^2 \tag{6}$$

where η_0 is the viscosity of the water (10^{-2} poise), l is the total length of the rotating cylinder, r is its radius (10^{-7} cm) and ω is the angular velocity. According to equation (6) ω increases with the progress of the unfolding reaction in view of the decrease in l.

$$l = (1{\cdot}7\times M_{(t)}/330)\ \text{Å} \tag{7}$$

where $M_{(t)}$ is the molecular weight of the still unfolded cylinder at a time t from the start of the reaction. Hence the time required to unwind a whole DNA molecule if the starting point is in the middle is given by

$$T = \int_0^{n/2}(2\pi/\omega)dn = \pi/3000\int_0^{M/2}(1/\omega)dM_{(t)} \tag{8}$$

where n is the number of the turns in the intact helix and M the molecular weight of DNA. Introducing ω and l from equations (6) and (7) and integrating we obtain

$$T \cong 0{\cdot}02\pi^2\eta r^2M^2 \cong 2\times10^{-17}M^2 \tag{9}$$

According to this model it takes only about 0·002 sec to unfold a molecule of molecular weight 10^7.

The process may however be slowed down considerably if some activation energy is involved. The upper limit for the time of unfolding must, according to our experiments, be less than one second but the actual value is probably much lower. We may take this to indicate that the activation energy cannot exceed a few kT.

According to the proposed model the rate of the unfolding reaction increases with its progress. Taking into account also that contact between the surface and the rotating cylinder is hindered because of hydrodynamic repulsion, the chances that the unfolding starts at more than one point are not very large. In such a case the reaction could proceed to completion only at the two points which are closest to the two ends of the molecule, leaving the middle of the molecule at least partly in its double helical form. This kind of imperfection in unfolding mechanism could be responsible for the fact that the area of native DNA (86 Å^2) at positive potentials is slightly lower than that of the denatured DNA (93 Å^2), even though the difference is close to the experimental error. According to the model suggested by Longuet-Higgins & Zimm (1960) the unfolding starts at the ends of the chain and proceeds by rotation of the two unfolded sections of the strands around their common axis.

Besides the lower hydrodynamic resistance of a rotating cylinder as compared with two random coiled strands rotating around a common axis, the model proposed in this work has the advantage that the unfolding can start at any point in the chain. Even taking into account the possible looseness of the structure at the end of the helix, the start of the unwinding at some randomly picked point in the chain seems

more probable. This is even more true if we assume, as Longuet-Higgins & Zimm did, that all the hydrogen bonds are broken before the unwinding starts.

It is understood that in different systems the unfolding forces may be of different kinds. One could suggest for instance that in a living cell the unfolding process may take place on the surface of an enzyme. As the unfolding is followed by synthesis of two new strands of a DNA molecule to match the two unwound single strands, a complex mechanism of supply of the necessary energy for the process has to be assumed. If an analogy from the behaviour on the very unbiological mercury surface can be drawn, one could speculate that the unfolding takes place at a positively charged protein surface whereas the synthesis takes place at a neutral or slightly negatively charged environment from which, after forming the proper hydrogen bonds, the new DNA molecule emerges.

In vitro more obvious sources for unwinding forces can be visualized. At extremely low ionic strengths the torque required to rotate the unwinding cylinders may be produced by the electrostatic repulsion between the charges of the two opposite strands. At elevated temperatures the thermal energy may supply the required torque.

The author wishes to thank Dr. A. Silberberg and Dr. D. Elson for their critical reading of the manuscript and for their valuable comments. He is also grateful to Professor P. Doty for the gift of salmon sperm DNA and to Dr. U. Littauer for the gift of *E. coli* RNA.

REFERENCES

Cox, R. A., Jones, A. S., Marsh, G. E. & Peacocke, R. A. (1956). *Biochim. biophys. Acta*, **21**, 576.
Cox, R. A. & Peacocke, R. A. (1956). *J. Chem. Soc.* 2499.
Delahay, P. & Trachtenberg, I. (1957). *J. Amer. Chem. Soc.* **79**, 2355.
Doty, P., Marmur, J., Eisner, J. & Schildkraut, C. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 461.
Eisenberg, H. (1957). *J. Polymer Sci.* **25**, 257.
Frumkin, A. N. & Melik-Gaikazyan, V. I. (1951). *C.R. Acad. Sci. U.S.S.R.* **77**, 855.
Gorodetskaya, A. & Frumkin, A. N. (1938). *C.R. Acad. Sci. U.R.S.S.* **18**, 639.
Grahame, D. C. (1947). *Chem. Rev.* **41**, 441.
Grahame, D. C. (1949). *J. Amer. Chem. Soc.* **71**, 2975.
Gulland, J. M., Jordan, D. O. & Taylor, H. F. W. (1947). *J. Chem. Soc.* 1131.
Hermans, J., Jr. & Hermans, J. J. (1959). *J. Phys. Chem.* **63**, 170.
Jacobson, B. (1955). *J. Amer. Chem. Soc.* **77**, 2919.
Koryta, J. (1953). *Collect. Czech. Chem. Com.* **18**, 206.
Kurland, C. G. (1960). *J. Mol. Biol.* **2**, 83.
Littauer, U. Z. & Eisenberg, H. (1959). *Biochim. biophys. Acta*, **32**, 321.
Longuet-Higgins, H. C. & Zimm, B. H. (1960). *J. Mol. Biol.* **2**, 1.
Melik-Gaikazyan, V. I. (1952). *J. Phys. Chem. U.S.S.R.* **26**, 560.
Miller, I. R. (1961). *Trans. Faraday Soc.*, in the press.
Miller, I. R. (1960). *J. Phys. Chem.* **64**, 1790.
Miller, I. R. & Grahame, D. C. (1956). *J. Amer. Chem. Soc.* **78**, 3577.
Miller, I. R. & Grahame, D. C. (1957). *J. Amer. Chem. Soc.* **79**, 3006.
Miller, I. R. & Grahame, D. C. (1961). *J. Colloid Sci.* **16**, 23.
Pouyet, J. (1952). *C.R. Acad. Sci., Paris*, **234**, 152.
Reichmann, M. E., Rice, S. A., Thomas, C. A. & Doty, P. (1954). *J. Amer. Chem. Soc.* **76**, 3407.
Sheraga, H. A. & Mandelkern, L. (1953). *J. Amer. Chem. Soc.* **75**, 179.
Warner, R. C. & Breslow, E. (1958). *Proc. 4th Int. Congr. Biochem., Vienna*, 157.
Watson, J. D. & Crick, F. H. C. (1953). *Nature*, **171**, 737.

J. Mol. Biol. (1961) **3**, 241–252

Factors Determining the Specificity of Hemoglobin Synthesized in a Cell-Free System

Hildegard Lamfrom

Division of Biology, California Institute of Technology, Pasadena, California, U.S.A.

(*Received 2 December 1960*)

In the cell-free system from reticulocytes, hemoglobin is synthesized which is characteristic of the cell species. The major objective of this investigation was to determine how the specificity in hemoglobin synthesis is attained. Cellular components—microsomes, pH 5 fraction and supernatant—were prepared from the reticulocytes of rabbits and sheep. Various combinations of these components from the two species were incubated under appropriate conditions. [^{14}C]valine and [^{14}C]leucine were used to identify the newly formed hemoglobins, which were separated chromatographically. A correlation could then be established between the type of hemoglobin synthesized and that of the cellular components.

The experiments present evidence that the microsomal fraction is highly specific and that the pH 5 fraction is not. They also indicate that there may, in addition, exist a factor in the supernatant fraction controlling specificity of hemoglobin synthesis. A speculative mechanism accounting for such dual specificity is discussed.

Further experiments suggest the presence of yet another factor in the reticulocyte supernatant fraction which seems to be particularly concerned with the release of hemoglobin from reticulocyte microsomes.

1. Introduction

The current theory of protein synthesis proposes a mechanism in which three steps are experimentally distinguishable: (1) the activation of amino acids; (2) the transfer of activated amino acids to soluble RNA; and (3) the transfer of these amino acids to the microsomes where they become linked by peptide bonds.

Of the very large number of conceivable arrangements of amino acids in peptide chains a cell produces only a limited number. In a rabbit reticulocyte, for instance, over 90% of the soluble protein synthesized is hemoglobin. This protein is not only different from all other rabbit proteins, but also differs from hemoglobin of other animals.

It was the objective of the present study to determine how the specificity in hemoglobin synthesis is attained. It has usually been assumed that the microsomes contain a specific "template" for protein synthesis (Littlefield, Keller, Gross & Zamecnik, 1955; Littlefield & Keller, 1957; Simkin & Work, 1957). However, experimental confirmation of such a hypothesis has been lacking. The reticulocyte cell-free system for synthesizing proteins has certain properties which enable a test to be made of this concept. With this system, as with the intact reticulocytes, almost all of the protein synthesized is hemoglobin; a large fraction of this newly synthesized hemoglobin is released from the microsomes; and it is possible to prepare cell-free systems from reticulocytes of different species which synthesize hemoglobin molecules separable by ion exchange chromatography. Therefore, the three components,

microsomes, pH 5 fraction, and supernatant were prepared from reticulocytes of different animal species (rabbit and sheep) and incubated in various combinations under appropriate conditions. Separation and identification of newly synthesized hemoglobin allowed a correlation to be established between its species type and that of the cellular components.

In the course of these experiments information was gathered on the release of protein from the microsomes which was controlled by a factor in the supernatant fraction. This process is separable from the three previously enumerated, and forms another experimentally distinguishable step in protein synthesis.

2. Materials and Methods

(a) *Reticulocytes*

Reticulocytes were obtained from two species of animals: rabbits and sheep. In rabbits reticulocytosis was induced by phenylhydrazine as described by Borsook, Deasy, Haagen-Smit, Keighley & Lowy (1952). The cells, often more than 95% of which were reticulocytes, were washed once with a Krebs-Henseleit solution containing 0·13 M-NaCl, $5{\cdot}2 \times 10^{-3}$ M-KCl and $7{\cdot}5 \times 10^{-3}$ M-$MgCl_2$ (NKM) and packed by centrifugation.

In sheep (type B) reticulocytosis was induced either with phenylhydrazine or by bleeding. With the phenylhydrazine technique an 85 lb sheep was injected subcutaneously on 7 consecutive days with 13·5 ml. of a 2·5% neutralized phenylhydrazine solution. With the second technique, approximately 500 ml. of blood was taken from an 80 lb sheep on each of 7 out of 9 days. In both instances blood for experimental use was removed when a reticulocytosis of 13 to 16% was reached, and the hemoglobin (Hb) had dropped to 3 to 4%. The cells were washed with NKM and stratified by centrifuging for 30 min at 3000 rev/min, concentrating reticulocytes in the top layer. The fraction removed contained approximately 50% reticulocytes. These were washed once in NKM and packed by centrifugation.

(b) *Preparation of cell fractions*

All operations were performed at 0° to 4°C, until the preparations were ready for incubation. The washed, packed reticulocytes were lysed with 4 vol. of 0·005 M-$MgCl_2$ and gently stirred for 10 min. After addition of 1 vol. of 1·5 M-sucrose containing 0·15 M-KCl, the mixture was centrifuged at 12,000 *g* for 10 min. The supernatant fraction (s_1) was centrifuged for 2 hr at 105,000 *g* in a Spinco preparative centrifuge to sediment the microsome fraction. The supernatant (s_2) was separated further into the pH 5 precipitate and the pH 5 supernatant fraction (sup.) by adjusting the pH to 5·2 with N-HAc. The precipitated protein was dissolved in 0·1 M-tris buffer (2-amino-2-hydroxymethyl-propane-1:3-diol) at pH 7·5, neutralized with 0·2 N-NaOH, centrifuged 2 hr at 105,000 *g* to remove contaminating microsomes,† and finally brought to a concentration of 12 to 15 mg protein per ml. (pH 5 fraction).

The supernatant was adjusted to pH 7·5 with N-NaOH. It was re-centrifuged 2 hr at 105,000 *g* but was always found free of sedimentable material.

The microsomes were washed by homogenizing with half the s_1 vol. of medium A (0·35 M-sucrose, 0·015 M-KCl, 0·035 M-$KHCO_3$ and 0·004 M-$MgCl_2$; Keller & Zamecnik, 1956) and centrifuged 2 hr at 105,000 *g*. The washed microsomes were suspended in medium A.

Guinea-pig liver cell fractions were prepared from 30% tissue homogenates in the same manner, except that the pH 5 fraction was precipitated twice.

(c) *Incubation conditions*

The conditions of incubation have been described previously (Schweet, Lamfrom & Allen, 1958) and are given for each experiment in the legends to the Tables. In addition

† Additional centrifugations up to 36 hr always yielded more sediment. Successive centrifugations sedimented material with progressively higher protein to RNA ratios, as determined spectrophotometrically.

to microsomes, pH 5 fraction, and supernatant, 1·5 ml. of the reaction mixture contained: 0·25 μmole guanosine triphosphate, 1 μmole adenosine triphosphate, 20 μmoles creatine phosphate, 190 μg creatine-kinase, 0·1 μmole glutamine, 0·05 ml. of a complete amino acid mixture minus the particular radioactive amino acid(s) employed, and 0·1 μmole of ^{14}C-labeled valine and/or leucine (uniformly labeled L-isomers obtained from Nuclear Chicago).

(d) *Resolution of incubation mixtures into microsome and soluble protein fractions*

Fragmentation of the microsomes occurred during the 1 hr incubation at 37°C, so that only 75% could be sedimented in 2 hr at 105,000 *g*. Therefore, microsomes were quantitatively precipitated, together with the pH 5 fraction, by adjusting the pH of the reaction mixture to 5·1. The precipitate was resuspended in medium A and will be referred to as "microsomes." Subsequent treatment of the microsomes with NaOH (described below), prior to the determination of radioactivity, splits the amino acid—S-RNA linkage. The supernatant from the pH 5·1 precipitation contained the "soluble proteins."

(e) *TCA precipitation and counting procedure*

In order to remove adsorbed free radioactive amino acids it was essential to include in the purification procedure 4 washes with the corresponding non-radioactive form of the amino acids. This was done as follows. Carrier protein was added to samples where necessary to give a total of 10 to 15 mg protein. The proteins were precipitated with 7% trichloroacetic acid (TCA), washed twice with 3·5% TCA containing 0·05 mole of the ^{12}C form of the amino acid (AA), reprecipitated, washed twice with 3·5% TCA containing the [^{12}C]AA, washed once with 3·5% TCA, and washed twice with acid acetone (0·5% HCl), twice with acetone and twice with ether. The dried precipitate was plated, weighed and counted on a Nuclear gas-flow counter with micromil end-window. Self absorption corrections to infinite thinness were made for all samples. Where cts/min/mg microsomes are reported, the estimation of microsome amount is based on absorbancy measurements (O.D.) at 259·5 mμ (mg microsomes/ml. = O.D. at 259·5 mμ $\times$ 0·089 $\times$ dilution factor).

(f) *Isolation of hemoglobin*

Soluble proteins from the fractionated incubation mixtures were dialysed for a total of 48 hr against three 1 l. lots of the developer used for subsequent chromatography, three 1 l. lots of developer containing 0·05 M-[^{12}C]valine and/or leucine and two 1·5 l. lots of developer without the amino acids. Carrier hemoglobin was dialysed separately against two changes of developer, and subsequently added to the experimental samples as described in the legends of the tables. Hemoglobin concentrations were based on spectrophotometric analysis at 541 mμ (mg Hb/ml. = O.D. at 541 $\times$ 1·14 $\times$ dilution factor) and 563 mμ (mg Hb/ml. = O.D. at 563 $\times$ 1·77 $\times$ dilution factor). After this dialysis the samples were still contaminated with small amounts of radioactive amino acids and a nondialysable, non-TCA-precipitable material of unknown nature. Some of this contamination was removed in the supernatant during a centrifugation at 105,000 *g* for 24 to 35 hr, which was introduced to concentrate the hemoglobin into 1 to 2 ml. The remainder of the contaminating material was separated from the major hemoglobin component upon chromatography.

Samples containing only rabbit hemoglobin were chromatographed at 4°C on Amberlite CG 50 (IRC-50) columns, using developer 1A (0·068 M-Na^+, pH 7·08; according to Allen, Schroeder & Balog, 1958). After elution of the minor components the column was warmed to room temperature (RT) to elute the rabbit major component.

For samples containing mixtures of sheep and rabbit hemoglobins developer 6D (0·05 M-Na^+, pH 6·73) was used. After elution of the minor components at 4°C the column was warmed to RT for elution of the sheep major component. The rabbit major component remained adsorbed at the top zone of the resin under these conditions, and was eluted with developer 1A at RT after transfer to a smaller column.

Optical densities at 280 mμ and 420 mμ were recorded for each fraction. The radioactivity of each fraction was determined on 1·5 ml. samples plated directly, and on samples subjected to TCA precipitation and cold amino acid washes as described above. For samples of the major hemoblogin components good agreement was obtained with the two methods. However, for the minor hemoglobin components good agreement was not obtained,

presumably because they were contaminated with adsorbed free [^{14}C] amino acids and other non-TCA precipitable material that also were eluted near the solvent front. In control experiments all adsorbed [^{14}C] amino acids could be removed from the minor hemoglobin components when the non-radioactive amino acid washes were included in the TCA treatment.

(g) *Fingerprints and radioautographs of hemoglobin*

Portions from the chromatographic peaks of rabbit and sheep major hemoglobin components were concentrated by centrifuging 24 hr at 105,000 *g* and dialysed against several changes of distilled water. Trypsin hydrolysis and fingerprinting were then carried out according to the procedure of Ingram (1958). Radioautographs were obtained after exposing the fingerprints to X-ray film for 1 to 8 months.

3. Experiments and Results

(a) *Rabbit reticulocyte—guinea-pig liver system*

Cellular fractions from guinea-pig liver (GP) and rabbit reticulocytes (R) were incubated in various combinations. The partition of incorporated [^{14}C]leucine between microsomes and soluble protein in these experiments is reported in Table 1.

Table 1

Distribution of incorporated [^{14}C] leucine in a system containing cellular components from guinea-pig liver and from rabbit reticulocytes

	Additions			Cts/min in	
	microsomes	pH 5 fraction	sup.	microsomes	total soluble protein
1.	GP	GP	GP	476	35
2.	GP	GP		476	27
3.	GP			153	25
4.		GP	GP	88	38
5.	GP	GP	R	648	36
6.	GP	R	R	647	51
7.	R	R	R	15,100	3,080
8.	R	R		16,200	649
9.	R		R	3,300	346
10.	R			1,500	105
11.		R	R	28	49
12.	R	GP	R	16,100	2,680
13.	R	GP	GP	16,900	824
14.	R	GP		9,730	485
15.	R		GP	5,100	241

Reaction mixtures were incubated for 30 min at 37°C and contained in a total volume of 1·5 ml. reagents as described under Methods. The following solutions were added as indicated. From GP: 0·5 ml. microsomes (9·25 mg ribonucleoprotein), 0·4 ml. pH 5 fraction (12 mg protein), 0·2 ml. sup. From R: 0·5 ml. microsomes (11·1 mg ribonucleoprotein), 0·4 ml. pH 5 fraction (6·8 mg protein), 0·2 ml. sup. The specific activity of [^{14}C]leucine was 5,300 cts/min/mμmole. Radioactivity is reported for TCA precipitated samples.

Comparing the all-rabbit system (line 7) with the all-guinea-pig system (line 1) one sees that the rabbit reticulocyte system is much more active than the guinea-pig liver system. Due to this low activity in the guinea-pig system nothing can be concluded about the amount of incorporation of labeled amino acids into soluble proteins. In the mixed experiments, the pH 5 fraction from rabbit reticulocytes and the pH 5 fraction from guinea-pig liver functioned equally well with the microsomes from both rabbit and guinea-pig (lines 5 and 6, 7 and 12). While rabbit supernatant appears to control the incorporation of radioactivity into the soluble protein it does not affect the incorporation into the microsomes. With no supernatant, 4% of the counts appeared in soluble protein (line 8). When supernatant was added 17% of the counts appeared in soluble protein (line 7). When rabbit supernatant was substituted for that from guinea-pig the radioactivity in the soluble protein was increased threefold, while the incorporation into the microsomal fraction remained the same (lines 12 and 13).

TABLE 2

Incorporation of [^{14}C]leucine into rabbit hemoglobin in a system containing cellular components from guinea-pig liver and from rabbit reticulocytes

	Additions			Cts/min in soluble protein				% total soluble cts in Hb major component
						Hb		
	microsomes	pH 5 fraction	sup.	put on column	recovered on column	minor component	major component	
1	R	R	R	2400	2760	169	2290	85
2	R	GP	R	1770	1550	98	1220	80
3	R	GP	GP	288	275	40	59	†
4	R			176	241	83	61	†

† low counts make consideration of ratios meaningless.

Conditions of incubation were identical with those given in Table 1. The soluble proteins were dialysed, concentrated and chromatographed as described under Methods. Carrier R hemoglobin was added to a final concentration of 10 mg. Radioactivity values were obtained from samples plated directly.

In order to show that counts incorporated into soluble protein represented amino acids incorporated into peptide bonds in hemoglobin, the following experiment was done. The soluble protein from an incubation of a mixture of rabbit microsomes, rabbit supernatant and either rabbit or guinea-pig pH 5 fraction was characterized by chromatography (Table 2). At least 80% of the incorporation was into rabbit major hemoglobin, regardless of the source of the pH 5 fraction. That incorporation was actually into true peptide linkage within the protein of the hemoglobin was confirmed by radioautography of the fingerprint from a trypsin digest of the rabbit major hemoglobin component. Several of the typical peptide spots of the fingerprint known to contain leucine were found to contain radioactivity.

The difference in total soluble counts reported for identical experiments in Tables 1 and 2 may be accounted for by probable proteolytic action occurring during dialysis. After that step the radioactivity could not be quantitatively sedimented together with carrier hemoglobin by centrifuging for 36 hr at 105,000 *g*. The missing counts could all be accounted for in the non-sedimentable fraction and proved to be non-precipitable by TCA. Good recovery was achieved in the chromatographic separation.

TABLE 3

Incorporation of [^{14}C]AA into microsome-bound and soluble protein in a mixed system from rabbit and sheep reticulocytes

	Additions			Experiment 1			Experiment 2	
				cts/min in			cts/min in	
	microsomes	pH 5 fraction	sup.	total sample	microsomes	soluble protein	microsomes	soluble protein
1.	R	R	R		14,000	3,700	69,400	65,400
2.	R				1,060	149	3,310	240
3.		R	R	28†			582	2,080
4.		R		38†				
5.			R	31†				
6.	R	R			16,000	649		
7.	R		R		3,330	346	66,600	60,200
8.	S	S	S		1,250	277	2,850	2,910
9.	S						608	135
10.		S	S	21†			52	124
11.		S		26†				
12.			S	35†				
13.	S		S				4,330	1,000
14.		R	S	25†			160	272
15.		S	R				60	96

† not significant.

Reaction mixtures were incubated for 60 min at 37°C with final concentrations of reagents as described under Methods. The reaction volume for experiment 1 was 1·5 ml. and for experiment 2, 6·0 ml. The specific activity of leucine was 3,600 cts/min/mμmole and of valine 5,100 cts/min/mμmole. The following solutions were added as indicated:

		Experiment 1	Experiment 2
Sheep:	microsomes	0·5 ml. = 6·6 mg ribonucleoprotein	1·3 ml. = 13·9 mg ribonucleoprotein
	pH 5 fraction	0·5 ml. = 4·42 mg protein	0·8 ml. = 12 mg protein
	sup.	0·2 ml.	2·6 ml.
Rabbit:	microsomes	0·5 ml. = 11·1 mg ribonucleoprotein	1·0 ml. = 22·2 mg ribonucleoprotein
	pH 5 fraction	0·5 ml. = 2·62 mg protein	0·8 ml. = 12 mg protein
	sup.	0·2 ml.	2·6 ml.

Radioactivity is reported for TCA precipitated samples.

(Reticulocytosis in the sheep was induced with phenylhydrazine in experiment 1 and by bleeding in experiment 2.)

(b) *Rabbit-sheep reticulocyte system*

Distribution of radioactivity between microsomes and soluble protein

The experiments described above for cell components from rabbit reticulocytes and from guinea-pig liver were repeated using cell components from rabbit reticulocytes and from sheep reticulocytes (S). In the second part of these experiments it

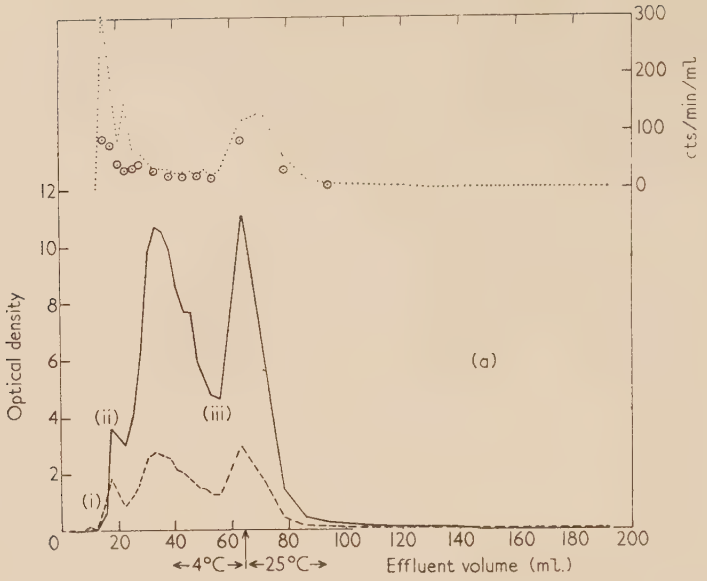
was therefore possible to characterize the soluble protein and ascertain whether the counts were in sheep or in rabbit Hb.

In Table 3 are reported results from two experiments in which the distribution of counts in microsomes and soluble protein was studied after incubation of components from sheep and rabbit reticulocytes in various combinations. The amount of supernatant for each incubation was 6 times greater in experiment 2 than in experiment 1. The rabbit supernatant fraction is somewhat contaminated with elements from the pH 5 fraction. When only 0·2 ml. supernatant is used per incubation, as in experiment 1, this is not noticeable. But when 1·2 ml. supernatant is employed as in experiment 2, the contamination becomes appreciable; the requirement for the pH 5 fraction is thereby alleviated (*cf.* lines 1 and 7 of experiment 2). The results are in all respects similar to those shown in Table 1. In agreement with these experiments, it was found that in the absence of both pH 5 fraction and supernatant the incorporation into microsomes and the percentage incorporation into soluble protein is much reduced (lines 1 and 2, 8 and 9). When only the supernatant was omitted the total incorporation into the microsomes remained the same, while incorporation into soluble protein was greatly diminished (lines 1 and 6). A similar effect of supernatant in mediating the release of counts into soluble protein is seen by comparing experiments 1 and 2 for lines 1, 7 and 8. By increasing the amount of supernatant sixfold the percentage of counts in soluble protein increased from 20 to 50.

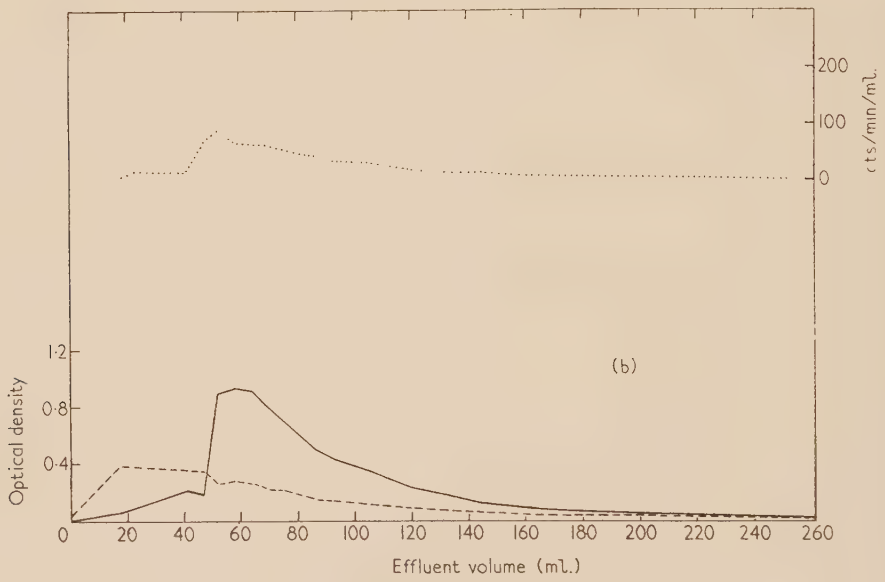
The second part of these experiments, which involves mixed rabbit-sheep systems, is complicated by two experimental obstacles. The rabbit system is much more efficient than the sheep system (lines 1 and 8), and the individual fractions are not pure; i.e. amino acid incorporation takes place in the absence of the microsomal fraction (lines 3 and 10) and it is unaffected by the omission of the pH 5 fraction (line 7). Most probably the pH 5 fraction of rabbit and sheep contain some contaminating microsomes in spite of re-centrifugation at 105,000 *g* for 2 hr, and the rabbit supernatant is contaminated with some pH 5 fraction.

Distribution of radioactivity between sheep and rabbit hemoglobin

The components from rabbit and sheep reticulocytes described in Table 3 were incubated and the soluble proteins analysed chromatographically. The results are presented in Table 4. Figure 1 gives an example of the elution diagram from a chromatographic column on which the data of Table 4 are based. Since the specific activity of the radioactivity is not constant throughout a peak, this indicates that sheep hemoglobin, as chromatographed, is composed of several molecular species. Radioautography of the fingerprints of the chromatographically isolated sheep and rabbit hemoglobins showed numerous of the typical peptide spots to be radioactive. This gives additional evidence that the counts incorporated into soluble protein are in hemoglobin. The data given in Table 4 demonstrate that the pH 5 fraction is not species-specific as judged by the distribution of radioactivity in the two hemoglobin species (lines 1 and 3, 5 and 7). The apparent effect of the rabbit pH 5 fraction in line 7 can be explained by the presence of contaminating microsomes, as shown in lines 3 and 14 of Table 3. That the rabbit supernatant is much richer in "releasing" factor than the sheep supernatant is reflected in the larger total amount of hemoglobin synthesized (*cf.* lines 1 and 4, 3 and 2, 8 and 5, 6 and 7). Moreover, even more sheep hemoglobin is obtained from sheep microsomes when rabbit supernatant supplies the "releasing" factor rather than sheep supernatant (*cf.* lines 6 and 8 to lines 5 and 7).



(a) Chromatogram of R and S non-heme component (i), R and S hemoglobin minor component (ii), and S hemoglobin major component (iii) on a 1×35 cm column of Amberlite CG 50 with developer 6D. Column temperature: 4°C (1 to 64 ml.) and 25°C (64 to 192 ml.).



(b) Chromatogram of R hemoglobin major component with developer 1A at 25°C, as described in text.

- - - - optical density at 280 mμ;
——— optical density at 420 mμ;
. . . . . cts/min/ml. for samples plated directly;
⊙ ⊙ ⊙ cts/min/ml. for samples counted after TCA precipitation.

FIG. 1. Elution diagram of mixed rabbit-sheep incubation system containing R microsomes, S pH 5 fraction and S supernatant.

TABLE 4

Incorporation of [^{14}C]AA into major Hb components in a mixed system from rabbit and sheep reticulocytes

	Additions			Cts/min in	Cts/min in Hb major component			% of total soluble cts/min
	microsomes	pH 5 fraction	sup.	microsomes	total in S and R	S	R	
								in R Hb
1.	R	R	R	74,700	86,800	648	86,100	99
2.	R	S	S	45,800	13,100	5,000	8,100	62
3.	R	S	R	58,000	82,500	2,380	80,000	97
4.	R	R	S	44,800	20,600	5,360	15,300	74
								in S Hb
5.	S	S	S	3,200	2,300	2,280	0	100
6.	S	R	R	9,000	13,000	10,400	2,590	80
7.	S	R	S	4,200	4,220	3,290	932	79
8.	S	S	R	3,200	7,770	5,340	2,430	69

Conditions of incubation were identical to those given for Table 3, experiment 2. The soluble proteins were dialysed, concentrated and chromatographed as described under Methods. Where necessary carrier hemoglobin of each species was added to a concn. of 10 mg. Approximately 60 mg hemoglobin was contributed by the supernatant fraction. Radioactivity of the microsomes is reported for TCA precipitated samples and of hemoglobin for samples plated directly.

When rabbit microsomes are incubated with sheep supernatant 26 to 38% of the total hemoglobin incorporation is into sheep hemoglobin (lines 4 and 2). That microsomal contamination in the supernatant is responsible for this effect can be ruled out by controls described in Table 3, lines 10 and 14. For a comparable incubation of sheep microsomes with rabbit supernatant 20 and 31% of the hemoglobin synthesized corresponds to the type of the supernatant (lines 6 and 8). Controls for this experiment show (Table 3, lines 3 and 15) that while microsomal contamination may play a role in the experiment using rabbit pH 5 fraction together with rabbit supernatant (line 6) this is unlikely when only rabbit supernatant is used (line 8). It may therefore be stated that in all cases a significant amount of incorporation into hemoglobin of both types was observed. However, it will be noted that always more than half and usually more than two thirds of the counts were found in the hemoglobin of the type corresponding to the microsomes used in the incubation. While this experiment presents further evidence that some specificity is associated with the microsomal fraction, it also indicates that there may, in addition, exist a factor in the soluble fraction controlling specificity of protein synthesis.

Note Added in Proof

A control experiment was performed to test for the possibility that the radioactivity found in the hemoglobin of the supernatant type might be due to the formation of a hybrid hemoglobin. One way this could conceivably have occurred would be by a chemical dissociation of hemoglobins into subunits when, at the end of the

experiment, the reaction mixture was acidified to pH 5·1 in order to separate the microsomal fraction from the soluble proteins.

To determine whether sheep and rabbit hemoglobins can undergo hybridization in acid solution, conditions were selected which were similar to those described by Vinograd & Hutchinson (1960) and which are optimal for the hybridization of rabbit and human A hemoglobins (Vinograd, J. & Morris, J., unpublished data). 25 mg of each ^{14}C rabbit and ^{12}C sheep carbon monoxyhemoglobin were combined and dialysed for 20 hr against 0·1 M-acetate buffer pH 4·65 containing 0·5% cysteine (Leif, R. C., personal communication), followed by 24 hr dialysis against the appropriate chromatographic developer at pH 7·0. The hemoglobins were then separated by chromatography and the eluates analysed for optical density and radioactivity.

Under these conditions only 2·6% of the total radioactivity of the combined major rabbit and sheep peaks appeared in the eluate with sheep hemoglobin. For comparison, if the ^{14}C rabbit and ^{12}C sheep hemoglobins were dialysed separately under the same conditions, and combined on the column, 1% of the radioactivity was transferred to the sheep hemoglobin fraction. In a simultaneously performed hybridization between the same preparation of ^{14}C rabbit and ^{12}C human A hemoglobin, 13·3% of the total counts in the major peaks were eluted with human hemoglobin.

It was therefore concluded that in the sheep-rabbit mixed incubation experiments reported in this paper the extent of labeling of the hemoglobin of the supernatant type (20 to 38%) is apparently not due to the particular experimental conditions employed subsequent to the incubation.

4. Discussion

Of all the cell free systems thus far studied, only in the reticulocyte and pea systems (Raacke, 1959; Webster, 1959) does a large fraction of the incorporated radioactive amino acids appear in soluble protein. In other systems (Hoagland, 1960) the synthetic process seems to stop with the assembly of protein or large peptides bound to the microsomes. Our experiments indicate the existence of a factor in the supernatant fraction of reticulocytes which is apparently concerned with the release of soluble protein material, i.e. hemoglobin, from the reticulocyte microsomes. Similar supernatant fractions from other sources, e.g. guinea-pig liver, do not appear to have this activity. The "releasing" factor from any of the reticulocytes tested (rabbit, sheep, duck, human [type A and type S]) functions interchangeably in all reticulocyte systems (Lamfrom, unpublished data). It is possible that the particular "releasing" factor in the reticulocyte supernatant is specific for hemoglobin synthesis as it is not active in other systems—e.g. radioactive proteins are not released from microsomes in the cell free systems from lymphocytes of immunized rabbits (Lamfrom, unpublished data).

When examining the specificity of the other cellular components the data indicate that the pH 5 fraction is neither species nor tissue specific. In the reticulocyte system this component derived from rabbit or human A or sickle cell reticulocytes can be freely interchanged and can also be replaced by that from guinea-pig liver. Likewise, the rabbit reticulocyte pH 5 fraction is effective in a system where amino acids are incorporated into guinea-pig microsomes.

The results presented also suggest that the microsomes play a role in determining the specificity of the protein synthesized. In all experiments the highest incorporation

of ^{14}C labeled amino acids occurred into the hemoglobin corresponding to the microsomal fraction. However, in every mixed system a significant amount of both types of hemoglobin was synthesized. The impurity of the cell fractions in the experiments reported can only be partly responsible for this phenomenon and it therefore appears that there may also exist an additional specificity in the supernatant fraction.

We can speculate about a mechanism for protein synthesis which would account for this kind of dual specificity. Such a scheme for protein synthesis has already been developed by Jacob (Riley, Pardee, Jacob & Monod, 1960; Jacob & Monod, 1961). Jacob's hypothesis postulates the existence of "messenger-RNA" (M-RNA), a special class of RNA molecules which are the primary gene products, and which encode the amino acid sequence of proteins. This M-RNA would be released into the cytoplasm and would participate in protein synthesis only while associated, temporarily, with microsomes. Such a scheme assumes that microsomes have a nonspecific structural function, while the M-RNA has the specific information for the synthesis of a particular protein. Perhaps each M-RNA is able to participate in the synthesis of only one protein molecule, and then must be replaced.

We imagine that in our experiments rabbit reticulocyte microsomes, as prepared from a cell lysate, are charged with M-RNA possessing the code for making rabbit hemoglobin. Additional M-RNA is also present in the supernatant fraction. When rabbit microsomes are incubated with rabbit supernatant they can, of course, be reloaded only with more rabbit M-RNA. But when rabbit microsomes are incubated with sheep supernatant, any additional M-RNA that becomes associated with empty or vacated microsomes could only be sheep M-RNA; thus sheep hemoglobin will also be synthesized.

Generalizing, then, one could expect that in mixed incubations where the microsomes and supernatant fractions are derived from two different species of reticulocytes, both types of hemoglobin will be synthesized. The hemoglobin corresponding to the microsomal type will be synthesized by M-RNA which remained associated with the microsomes during the isolation procedure; the hemoglobin of the supernatant type will be synthesized by new M-RNA supplied by this cell fraction.

The results of the rabbit-sheep mixed incubation experiments reported here do not contradict the "messenger RNA" concept; it is hoped that experiments now in progress will show whether such a scheme is actually valid.

It is a pleasure to thank Dr. H. Borsook for his interest and kind advice and Drs R. T. Jones, A. Miller, C. Steinberg and H. Temin for helpful discussions. I am indebted to Mrs. Miriam Wright for her expert technical assistance. This work was supported in part by a grant (H-1624) from the United States Public Health Service.

REFERENCES

Allen, D. W., Schroeder, W. A. & Balog, J. (1958). *J. Amer. Chem. Soc.* **80**, 1628.
Borsook, H., Deasy, C. L., Haagen-Smit, A. J., Keighley, G. & Lowy, P. H. (1952). *J. Biol. Chem.* **196**, 669.
Hoagland, M. B. (1960). In *The Nucleic Acids*, Vol. 3, ed. by E. Chargaff & J. N. Davidson, p. 349. New York: Academic Press.
Ingram, V. M. (1958). *Biochim. biophys. Acta*, **28**, 539.
Jacob, F. & Monod, J. (1961). *J. Mol. Biol.* **3**, 318.
Keller, E. B. & Zamecnik, P. C. (1956). *J. Biol. Chem.* **221**, 45.

Littlefield, J. W., Keller, E. B., Gross, J. & Zamecnik, P. C. (1955). *J. Biol. Chem.* **217**, 111.
Littlefield, J. W. & Keller, E. B. (1957). *J. Biol. Chem.* **224**, 13.
Raacke, I. D. (1959). *Biochim. biophys. Acta*, **34**, 1.
Riley, M., Pardee, A. B., Jacob, F. & Monod, J. (1960). *J. Mol. Biol.* **2**, 216.
Schweet, R., Lamfrom, H. & Allen, E. (1958). *Proc. Nat. Acad. Sci., Wash.* **44**, 1029.
Simkin, J. L. & Work, T. S. (1957). *Biochem. J.* **65**, 307.
Vinograd, J. & Hutchinson, W. D. (1960). *Nature*, **197**, 216.
Webster, G. E. (1959). *Fed. Proc.* **18**, 1379.

J. Mol. Biol. (1961) **3**, 253–256

Sur un nouveau Mécanisme d'Oxydo-réduction, appliqué à la Réaction d'Épimérisation: Uridinediphosphogalactose⇌Uridinediphosphoglucose†

Huguette de Robichon-Szulmajster
Laboratoire de Génétique Physiologique du C.N.R.S., Gif-sur-Yvette (S.O.), France

(*Received 3 December 1960*)

The epimerization reaction uridine diphosphogalactose ⇌ uridine diphosphoglucose is explained by a new theoretical mechanism of oxido-reduction, based on two hydrogen acceptors uracil and diphosphopyridine nucleotide, which is allowed by the molecular configuration of the substrates.

Au cours de la réaction d'épimérisation, UDPGal ⇌ UDPG,‡ catalysée par l'UDPGal-4-épimérase, les positions respectives occupées par l'—H et l'—OH du quatrième carbone se trouvent inversées (Leloir, 1951). Il a été montré que l'UDPGal-4-épimérase de foie de mammifère requiert le DPN et est inhibée par le DPNH (Maxwell, 1956), tandis que la même enzyme, extraite à partir de la levure, ne présente ni exigence pour le DPN, ni inhibition par le DPNH. L'enzyme de levure possède, par contre, une fluorescence caractéristique (Maxwell, de Robichon-Szulmajster & Kalckar, 1958), due à la présence de DPN lié à l'enzyme (Maxwell & de Robichon-Szulmajster, 1961). Toutefois, le mécanisme de la réaction reste obscur.

La construction d'un modèle moléculaire de l'UDPG nous a permis de constater que cette molécule peut exister, sans empêchement stérique, sous une forme repliée du type de celle proposée par Szent-Györgyi pour l'ATP (Szent-Györgyi, 1957). Sous cette forme (voir Fig. 1(b)§) le noyau de l'uracile, et particulièrement le groupement —CO—NH—CO— de ce noyau se trouve non seulement très proche, mais, de plus, symétriquement disposé de part et d'autre, au-dessus de la position 4 de l'hexose.

Il nous a semblé que cette configuration était en faveur de l'hypothèse avancée par Kalckar (1958), tendant à attribuer au noyau de l'uracile un rôle au moins partiel, au cours de la réaction d'épimérisation.

Afin de rendre compte, à la fois, de la stéréospécificité de l'enzyme (non-incorporation de tritium dans la molécule de l'hexose à partir d'eau tritiée ou de DPN tritié (Maxwell, 1956; Maxwell *et al.*, 1958; Kowalsky & Koshland, 1956)) et de la structure supposée de la molécule, nous proposons un schéma d'oxydo-réduction à double catalyse, dans lequel les deux catalyseurs, DPN et uracile, sont capables d'accepter conjointement, chacun un atome d'hydrogène provenant du substrat. Il est implicitement admis que, de même que dans le mécanisme classique de l'oxydo-réduction, l'H transféré au DPN sera toujours celui qui est lié au carbone (4), dans la molécule de l'hexose, tandis que l'H transféré à l'uracile proviendra toujours de l'hydroxyle de ce même carbone.

† Ce travail a bénéficié de l'aide de la Fondation Rockefeller.

‡ Abréviations: UDPGal, uridinediphosphogalactose; UDPG, uridinediphosphoglucose; DPN, DPNH, diphosphopyridine nucléotide oxydé ou réduit; ATP, adénosine triphosphate.

§ Afin de le différencier, sur le modèle, l'un des atomes d'hydrogène de l'hexose (celui du groupement —OH du 4ème carbone) a été remplacé par un atome de fluor.

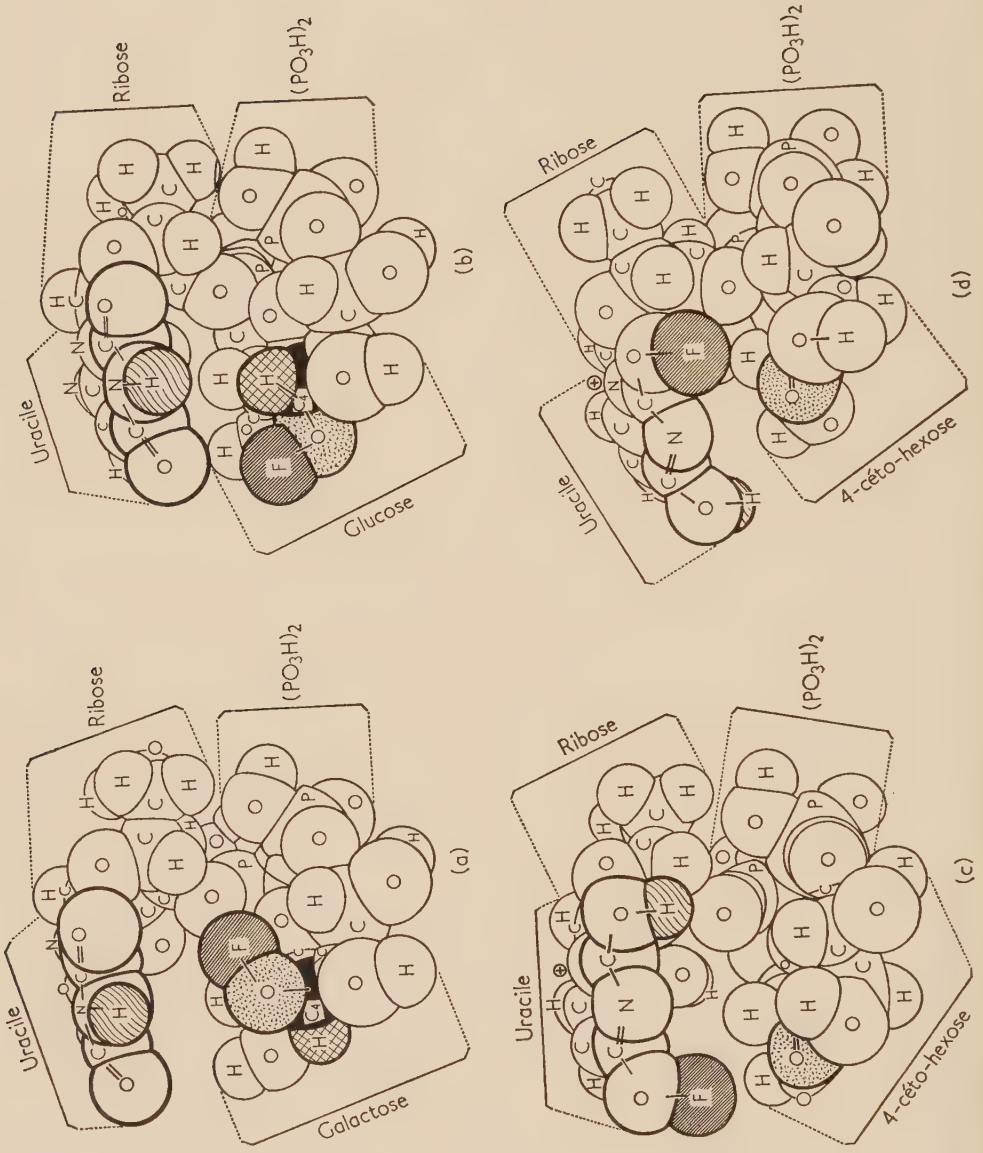

Uracile
Ribose
$(PO_3H)_2$
Glucose
(b)
Uracile
Ribose
$(PO_3H)_2$
4-céto-hexose
(d)
Uracile
Ribose
$(PO_3H)_2$
Galactose
(a)
Uracile
Ribose
$(PO_3H)_2$
4-céto-hexose
(c)

FIG. 1

Le schéma proposé est représenté dans la Fig. 2. On peut rendre compte des phases successives de la réaction de la manière suivante:

(1) 2 atomes d'hydrogène sont enlevés simultanément au 4ème carbone de l'hexose qui devient un 4-céto-hexose, intermédiaire hypothétique (voir aussi Fig. 1(c)).

(2) 1 atome d'hydrogène est fixé sur le DPN en position para;

1 atome d'hydrogène est fixé sur l'une des positions cétone de l'uracile. Selon que l'hexose de départ est le glucose (Fig. 1(b)) ou le galactose (Fig. 1(a)) l'une des deux positions sera plus proche que l'autre du groupement C_4—OH de l'hexose (Fig. 1(c) ou 1(d), respectivement).

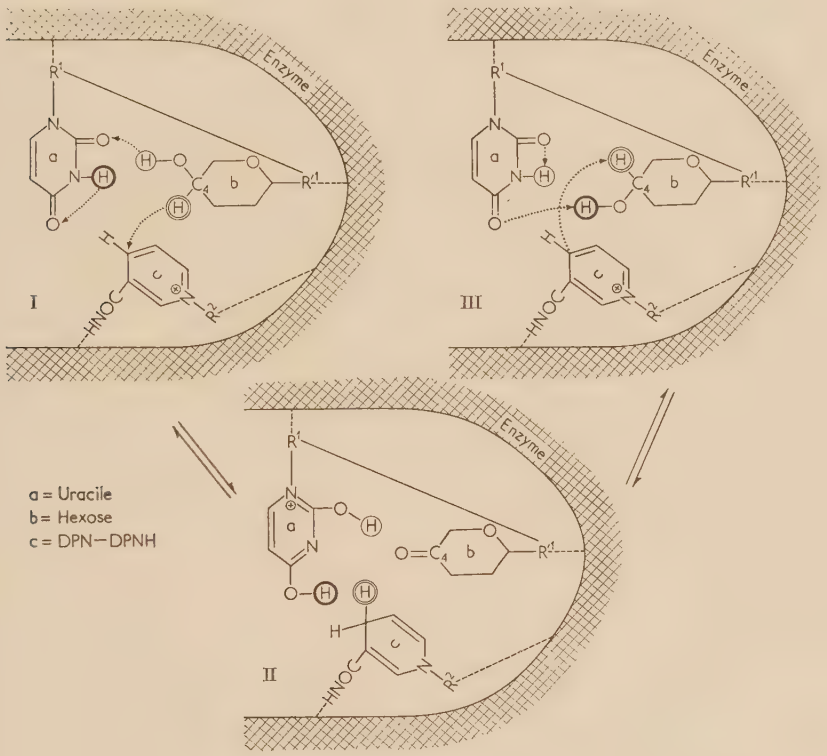


Fig. 2. Schéma de l'épimérisation UDPGal ⇌ UDPG.

(3) Simultanément, les 2 noyaux sont amenés au remaniement des doubles liaisons. En conséquence:

(*a*) sur le noyau de l'uracile, l'H lié à l'azote migre sur la position cétone restée libre (Fig. 1(c) et 1(d)). De ce fait, l'azote en position 3 devient $-\overset{+}{\underset{\|}{N}}-$ et perd un électron;

(*b*) sur le noyau du DPN, le $-\overset{+}{\underset{\|}{N}}-$ perd sa charge, en acquérant un électron.

(4) Le groupement —CO—NH—CO— de l'uracile étant devenu un diénol, $-\underset{|}{C}=N=\underset{|}{C}-$ dont les deux —OH sont situés à distance sensiblement égale du

OH OH

groupement cétonique en position (4) de l'hexose, le retour d'un atome d'hydrogène peut avoir lieu à partir de l'un ou de l'autre de ces hydroxyles, provoquant, ou non, l'inversion.

(5) Simultanément, les deux noyaux remanient à nouveau leurs liaisons: le DPNH retourne à la forme pyridinium$^+$ et perd un électron, tandis que l'uracile retourne à la forme cétonique, ce qui ramène sur l'azote l'H du second hydroxyle, et supprime la charge du noyau par acquisition d'un électron.

Au cours de ce processus d'oxydo-réduction d'un nouveau type, les deux atomes d'hydrogène enlevés au substrat seraient donc fixés sur deux noyaux différents, l'uracile et le DPN. Ceci est rendu possible par un phénomène de compensation des charges entre les deux noyaux.

Dans ce mécanisme, l'enzyme qui, on le sait, est capable de fixer le DPN (Maxwell & de Robichon-Szulmajster, 1961) provoquerait également la formation d'un complexe avec le substrat et favoriserait, de plus, la forme repliée de la molécule nucléotidique, telle qu'elle est représentée dans la Fig. 1, mettant ainsi en contact les trois composantes du système: hexose-uracile-DPN.

Le rôle qui, dans ce mécanisme, incombe à l'uracile est d'être à la fois et de façon réversible, accepteur d'hydrogène et donateur d'électron. Il semble hautement probable que, si tel est son rôle, le fait que l'uracile fasse partie de la molécule dans laquelle l'hexose est intégré et qu'il soit, par conséquent, placé dans les conditions stériques les plus favorables, doive lui conférer une grande spécificité dans la catalyse de la réaction d'épimérisation.

L'hypothèse exposée ci-dessus semble à l'heure actuelle difficile à mettre à l'épreuve puisque, d'une part, les atomes d'hydrogène liés à l'oxygène sont très rapidement échangeables (Vennesland & Westheimer, 1953) et que, d'autre part, toutes les tentatives tendant à capter un 4-céto-hexose intermédiaire ont été négatives (Maxwell, 1956; Maxwell *et al.*, 1958). Cet échec pourrait cependant s'expliquer par empêchement stérique, le réactif ne pouvant avoir accès au groupement cétonique de la forme intermédiaire de l'hexose (voir schéma).

Nous espérons que ces considérations théoriques sur un mécanisme d'oxydo-réduction basé sur une double catalyse pourront être applicables à d'autres réactions.

Nos remerciements vont à M. le Pr. H. M. Kalckar, avec qui de nombreuses discussions (au cours du séjour de l'auteur dans son laboratoire) ont permis la conception de ce mécanisme, à M. le Pr. B. Pullman, pour d'utiles discussions concernant la validité théorique du schéma proposé.

REFERENCES

Kalckar, H. M. (1958). *Advanc. Enzymol.* **20**, 111.
Kowalsky, A. & Koshland, D. E., Jr. (1956). *Biochim. biophys. Acta*, **22**, 575.
Leloir, L. F. (1951). *Arch. Biochem. Biophys.* **33**, 186.
Maxwell, E. S. (1956). *J. Amer. Chem. Soc.* **78**, 1074.
Maxwell, E. S., de Robichon-Szulmajster, H. & Kalckar, H. M. (1958). *Arch. Biochem. Biophys.* **78**, 407.
Maxwell, E. S. & de Robichon-Szulmajster, H. (1960). *J. Biol. Chem.*, **235**, 308.
Szent-Györgyi, A. (1957). *Bioenergetics.* New York: Academic Press.
Vennesland, B. & Westheimer, F. H. (1954). In *Symposium on the Mechanism of Enzyme Action*, ed. by W. D. McElroy & B. Glass, p 357. Baltimore: Johns Hopkins Press.

J. Mol. Biol. (1961) **3**, 257–262

The Polypeptide Chains of Haemoglobin-A_2 and Haemoglobin-G_2

E. R. Huehns† and E. M. Shooter

Department of Biochemistry, University College, London W.C.1, England

(*Received 11 December 1960*)

Dissociation and recombination experiments using Hb-A_2 with Hb-H and Hb-G_{Br} are described. Significant amounts of hybrid haemoglobins only appeared when haemoglobin mixtures were exposed to pH 4·7 for at least 72 hr. With Hb-H, Hb-A was formed and with Hb-G_{Br}, which has abnormal α_2^G subunits, a new zone slower than Hb-A_2 on starch gel electrophoresis appeared. No new haemoglobin species were detected when Hb-A_2, Hb-G_{Br} or Hb-H by themselves or a mixture of Hb-A_2 and Hb-S were dissociated and recombined in parallel experiments. From these results it is concluded that Hb-A_2 has the composition $\alpha_2^A\delta_2^{A_2}$. The genetic variants of Hb-A_2 fall into two groups, one independent of Hb-A and the other linked to α-chain variants of Hb-A. Hb-G_2, ($\alpha_2^G\delta_2^{A_2}$), which is found in conjunction with Hb-Gα, is of the latter type. Hb-G_2 occurs with Hb-A_2 in individuals that carry both Hb-A and Hb-G_α, and is the sole minor component in the individual who carries only Hb-G_α. Further it is suggested that any individual carrying an α-chain variant of Hb-A would also carry the corresponding minor component, namely $\alpha_2^X\delta_2^{A_2}$, and that therefore the synthesis of the α_2 subunits of Hb-A and Hb-A_2 are controlled at the same genetic locus. Synthesis of the $\delta_2^{A_2}$ subunit is controlled by a separate locus, the δ locus, variation of which it is suggested gives rise to the Hb-A independent minor components such as Hb-A_2' and Hb-B_2.

1. Introduction

Haemoglobin A_2 (Hb-A_2), the minor haemoglobin present in normal adult haemolysates, was first described by Kunkel & Wallenius (1955). At alkaline pH Hb-A_2 has an electrophoretic mobility on starch block, on paper, or in starch gel which is similar to that of haemoglobin E but less than that of haemoglobin A (Hb-A), the major haemoglobin of normal adult haemolysates. In normal adults Hb-A_2 accounts for approximately 2·5% of the total haemoglobin. The proportion of Hb-A_2 is often increased in thalassaemia, but it does not exceed 15% of the total haemoglobin, and this property serves to distinguish it from Hb-E which always occurs in excess of 20% of the total haemoglobin (Aksoy, Lehmann & Eng, 1957). Hb-A_2 can also be differentiated from Hb-A, but not from Hb-E, by chromatography on carboxymethylcellulose (Huisman, Martis & Dozy, 1958).

A number of other minor haemoglobins have been reported and these may be divided into two groups. The first group comprises those minor haemoglobins which occur together with Hb-A and the second those which occur only with variants of Hb-A. The two examples of the first group so far described are the minor haemoglobins Hb-A_2' (Ceppellini, Kunkel & Dunn, 1957) and Hb-B_2 (Ceppellini, 1959), and the

† Beit Memorial Fellow

study of a number of families in which they appear suggests that these haemoglobins are the product of a rare allele at the Hb-A_2 locus. The first minor haemoglobin of the second group, Hb-G_2, was found in an individual with Hb-G disease (Shooter, Skinner, Garlick & Barnicot, 1960). The haemolysate from this individual contained Hb-G_{Ib} and Hb-G_2, but neither Hb-A nor Hb-A_2. Hb-G_2 migrated at alkaline pH even slower than Hb-A_2. Gammack, Huehns, Shooter & Gerald (1960) pointed out that Hb-G_2 differs in charge from Hb-A_2 by about the same amount as Hb-G_{Ib} from Hb-A and therefore suggested that Hb-G_2 is the naturally occurring hybrid of Hb-G_{Ib} and Hb-A_2 with the composition $\alpha_2^G\delta_2^{A_2}$.†

This hypothesis implies not only that Hb-G_{Ib} has abnormal α_2 subunits, but also that Hb-A_2 consists of α_2^A subunits identical to those of Hb-A and a second subunit, $\delta_2^{A_2}$, characteristic of Hb-A_2, its molecular composition therefore being $\alpha_2^A\delta_2^{A_2}$. In support of this hypothesis Gammack *et al.* (1960) were able to show by means of hybridisation experiments that Hb-G_{Ib} has abnormal α_2^G and normal β_2^A subunits. The results of further hydridisation experiments, which are reported here, have now confirmed that Hb-A_2 has the molecular composition $\alpha_2^A\delta_2^{A_2}$.

2. Materials and Methods

Haemoglobin solutions. These were prepared by the method of Singer, Chernoff & Singer (1951) from fresh venous blood. All the experiments were made with carbon monoxy-haemoglobins. Solutions were stored under CO and dialyses were carried out against the appropriate CO-saturated buffer. If methaemoglobin was present this was reduced by adding the minimum amount of solid sodium dithionite. The reduced haemoglobin was subsequently converted to the CO derivative with carbon monoxide. This procedure was only carried out immediately prior to dialysis.

Isolation of individual haemoglobins. Haemoglobin-$G_{Bristol}$ (Hb-G_{Br}) (Raper, Gammack, Huehns & Shooter, 1960), which is essentially similar to Hb-G_{Ib} in that it also has abnormal α_2 subunits, was used in this investigation. Hb-G_{Br} was obtained from the haemolysate of the individual carrying the four haemoglobins A, G_{Br}, C and G_{Br}/C, and Hb-H from the cases previously reported by Bingle, Huehns & Prankerd (1958). Hb-A_2 was prepared either from the haemolysates of normal adults or from a case of thalassaemia (Case 1, Bailey & Prankerd, 1958) in whom Hb-A_2 amounts to 8% of the total haemoglobin.

Hb-G_{Br} and Hb-H were isolated from the appropriate haemolysate by starch block electrophoresis (Kunkel, 1954) following the procedure previously outlined (Gammack *et al.*, 1960). The dilute eluted haemoglobin solutions were concentrated by the method devised by Everall & Wright (1958) with modifications. One end of a 50 cm length of $\frac{1}{4}$ in dialysis tube was held in a hole of a rubber bung by a tapered piece of glass tubing. The other end was knotted twice and hung inside a 5 l. Büchner flask. This flask contained 2 l. of ice-cold 0·2 M-NaCl and the dialysis tubing dipped into this solution. Haemoglobin solution was pipetted into the dialysis sac through the glass tubing and the flask evacuated on a water pump to a pressure not less than 200 mm Hg and then kept in the cold. Using this method 20 to 30 ml. of dilute haemoglobin solution was concentrated to 0·5 ml. with a final concentration between 3 and 10 g/100 ml. in about 16 hr. The final preparations of both haemoglobins contained small amounts of Hb-A. The Hb-H preparation presumably also contained small amounts of Hb-"Bart's".

Hb-A_2 solutions prepared from a normal adult haemolysate by this procedure contained at least the same concentration of non-haem proteins as of Hb-A_2. Hb-A_2, free from non-haem protein, was prepared by chromatography followed by starch block electrophoresis. The chromatographic technique used was based on that described by Huisman *et al.* (1958), who found that Hb-A is eluted by 0·01 M-phosphate buffer from carboxymethylcellulose

† The non-α_2 subunit of Hb-A_2 is now commonly called the $\delta_2^{A_2}$ subunit.

(CMC) at pH 7·49, whereas Hb-A_2 is eluted at pH 7·69. The haemolysate (containing 8% of Hb-A_2) was diluted to a haemoglobin concentration of approximately 0·5 g/100 ml. and a total volume of 250 ml. with phosphate buffer (0·01 M-NaH_2PO_4 whose pH was accurately adjusted to 7·50 at 25°C with 10% NaOH) and then dialysed against 3 l. of the same phosphate buffer in the cold. The dialysed haemoglobin solution was applied to a $90 \times 0{\cdot}9$ cm column packed with CMC (Whatman CM 80 powder) which had previously been equilibrated by allowing 2 l. of the pH 7·50 buffer to flow through it. The flow rate of the column during the addition of the haemoglobin solution and subsequent operations was 10 ml./hr. After applying the haemoglobin, 250 ml. of the phosphate buffer was allowed to pass through the column. This buffer eluted the non-haem proteins and the major part of the Hb-A present. The remaining haemoglobin in the column was eluted with a second phosphate buffer of pH 8·0 (0·04 M-$Na_2HPO_4.12H_2O$ whose pH was adjusted to pH 8·0 with syrupy phosphoric acid, B.P.) and the eluate of about 100 ml. concentrated in the usual manner. At this stage the solution contained approximately 70% Hb-A_2 and 30% Hb-A; no non-haem proteins could be detected by analysis in starch gel. This concentrated solution was subjected to starch block electrophoresis and two well separated zones corresponding to Hb-A_2 and Hb-A were observed. The Hb-A_2 zone was cut out, the haemoglobin eluted and the solution concentrated as described above. Analysis in starch gel of the final preparation showed only one zone migrating in the position of Hb-A_2. About 80% yield of Hb-A_2 was obtained.

Starch gel electrophoresis followed the procedures already outlined (Gammack *et al.*, 1960).

Hybridisation experiments were made as previously described but with minor modifications. The total haemoglobin concentration of the initial mixture of the two appropriate haemoglobins was adjusted to between 3 and 4 g/100 ml. and after dividing into two equal volumes these solutions were then dialysed directly against either a sodium acetate buffer of pH 4·7 and *I* 0·10 or, for the control, phosphate buffer of pH 6·8 and *I* 0·02. The time of contact with the acetate buffer (or with the phosphate buffer for the control) was also extended beyond 4 hr.

3. Results

(a) *Hybridisation of Hb-A_2 with Hb-G_{Br}*

Preliminary experiments showed that Hb-A_2 did not form consistently detectable amounts of hybrid haemoglobins unless the time for which the haemoglobin mixture was maintained at pH 4·7 was greater than 48 hr and preferably of the order of 90 hr (Plate I). It was then possible to show that when mixtures of Hb-A_2 from the individual with thalassaemia and of Hb-G_{Br} were exposed to pH 4·7 for that time and subsequently dialysed to pH 6·8 a new haemoglobin zone migrating behind Hb-A_2 and in the position of the Hb-G_{Br}/C marker was formed, and some intensification of the zone in the position of Hb-A occurred (Plate II). In control experiments, where separate solutions of Hb-A_2 and Hb-G_{Br} were treated in the same way, no new haemoglobin zones appeared (Plate III). In contrast to the experiment with Hb-G_{Br} no change in composition was observed when a mixture of Hb-A_2 and Hb-S was exposed to pH 4·7 for 92 hr with subsequent dialysis to pH 6·8 (Plate II, (iv) and (v)).

A hybridisation experiment with Hb-A_2 from a normal adult gave a result essentially similar to those described above (Plate IV). In this instance, however, the quantity of Hb-A_2 relative to Hb-G_{Br} was very small and thus correspondingly less hybrid was formed.

(b) *Hybridisation of Hb-A_2 with Hb-H*

Dialysis of mixtures of Hb-A_2 (from the individual with thalassaemia) and Hb-H at pH 4·7 for 72 hr followed by dialysis to pH 6·8 produced significant amounts of a

haemoglobin migrating in the position of Hb-A (Plate V). It will be noted that the migration rates of this new zone and Hb-A are more nearly equal in the analysis of the experiment with the lower concentration of Hb-A_2 (Plate V(ii)). At the higher concentration of Hb-A_2 (Plate V(iv)) the movement of the new zone is slightly decreased by the presence of the more concentrated Hb-A_2 zone migrating towards the opposite electrode. Both control mixtures show contaminating traces of Hb-A from the isolated Hb-H specimen. The zone migrating just behind Hb-H in the two controls is probably Hb-"Bart's." In a parallel control experiment, where only Hb-H was treated with acetate, no new haemoglobin zones were formed. In all these experiments some precipitation occurred and only very small amounts of Hb-H could be detected when the pH was brought back to 6·8.

4. Discussion

Variants of Hb-A form hybrid haemoglobins readily after standing at pH 4·7 for 4 hr in the cold (Itano & Robinson, 1959; Gammack, Huehns, Lehmann & Shooter, 1961). In contrast, comparable amounts of hybrid formation were only obtained in experiments with Hb-A_2 when the time of dialysis against the pH 4·7 acetate buffer was 70 hr or more. It is apparent, therefore, that Hb-A_2 only exchanges its subunits relatively slowly. An attempt was made to displace the exchange reaction in favour of hybrid formation by increasing the proportion of the major haemoglobin variant, either Hb-H or Hb-G_{Br}, relative to that of Hb-A_2. However, even when there was ten times as much of the major haemoglobin present as Hb-A_2, more of the latter reformed than took part in hybrid formation.

The appearance of Hb-A in the hybridisation of Hb-A_2 with Hb-H indicates that Hb-A_2 has α_2 subunits similar to those of Hb-A. Although the reaction can probably be described in the usual manner (Jones, Schroeder & Vinograd, 1959)

$$\underset{\text{Hb-}A_2}{\alpha_2^A\delta_2^{A_2}} + \underset{\text{Hb-H}}{\beta_4^A} \longrightarrow \underset{\text{Hb-A}}{\alpha_2^A\beta_2^A} + \underset{\text{Hb-}A_2}{\alpha_2^A\delta_2^{A_2}} + \beta_2^A + \delta_2^{A_2}$$

the exact stoichiometry cannot be specified for the experiments reported here. The fate of the excess β_2^A or $\delta_2^{A_2}$ subunits has not been investigated. (On the same basis any Hb-"Bart's" present would form Hb-F with Hb-A_2. However, Hb-A and Hb-F are not resolved under the conditions of this analysis.) The formation of Hb-A as well as a second hybrid haemoglobin migrating slower than Hb-A_2 in the hybridisation of Hb-A_2 with Hb-G_{Br} suggests that Hb-A_2 consists of an α_2^A subunit joined to a non-α_2 subunit, the $\delta_2^{A_2}$ subunit, which differs from the β_2 subunit of Hb-A. It is clear from the starch gel analysis (Plate II) that the difference in net charge between the slow-moving hybrid and Hb-A_2 is the same as that between Hb-G_{Br}/C and Hb-C (the two slower-moving zones in the marker). Since the only structural difference between Hb-G_{Br}/C and Hb-C is the replacement of the α_2^A subunit in Hb-C by an $\alpha_2^{G_{Br}}$ subunit in Hb-G_{Br}/C it follows that the probable composition of the hybrid haemoglobin is $\alpha_2^G\delta_2^{A_2}$. The hybridisation experiment may therefore be represented by

$$\underset{\text{Hb-}G_{Br}}{\alpha_2^{G_{Br}}\beta_2^A} + \underset{\text{Hb-}A_2}{\alpha_2^A\delta_2^{A_2}} \longrightarrow \underset{\text{Hb-A}}{\alpha_2^A\beta_2^A} + \underset{\text{Hb-}G_{Br}}{\alpha_2^{G_{Br}}\beta_2^A} + \underset{\text{Hb-}A_2}{\alpha_2^A\delta_2^{A_2}} + \underset{\text{Hb-}G_{Br}/A_2}{\alpha_2^{G_{Br}}\delta_2^{A_2}}$$

Charge difference from Hb-A	+ 2	+ 4	0	+ 2	+ 4	+ 6

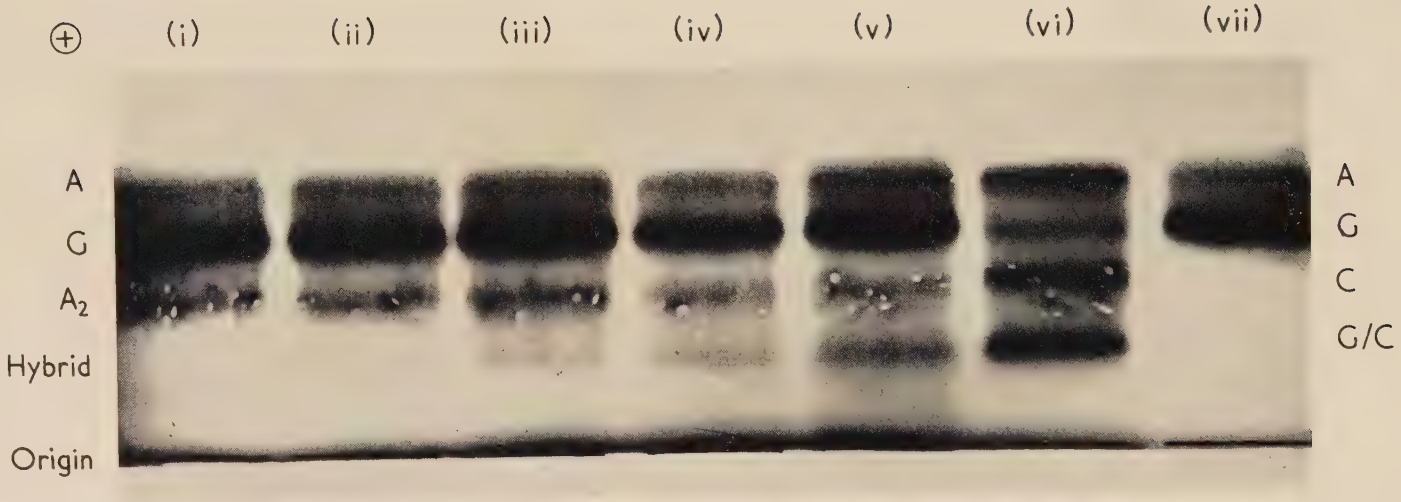


PLATE I. Hybridisation of haemoglobin A_2 and G_{Br}. Starch gel analysis in the discontinuous system. Migration time 2 hr, *o*-dianisidine stain. Mixture at pH 4·7 for (i) 0 hr (ii) 12 hr (iii) 24 hr (iv) 48 hr (v) 96 hr with subsequent dialysis to pH 6·8. (vi) Hb A, G_{Br}, C and G_{Br}/C marker. (vii) Hb-G_{Br} kept at pH 4·7 for 96 hr with subsequent dialysis to pH 6·8.

[*To face page* 260

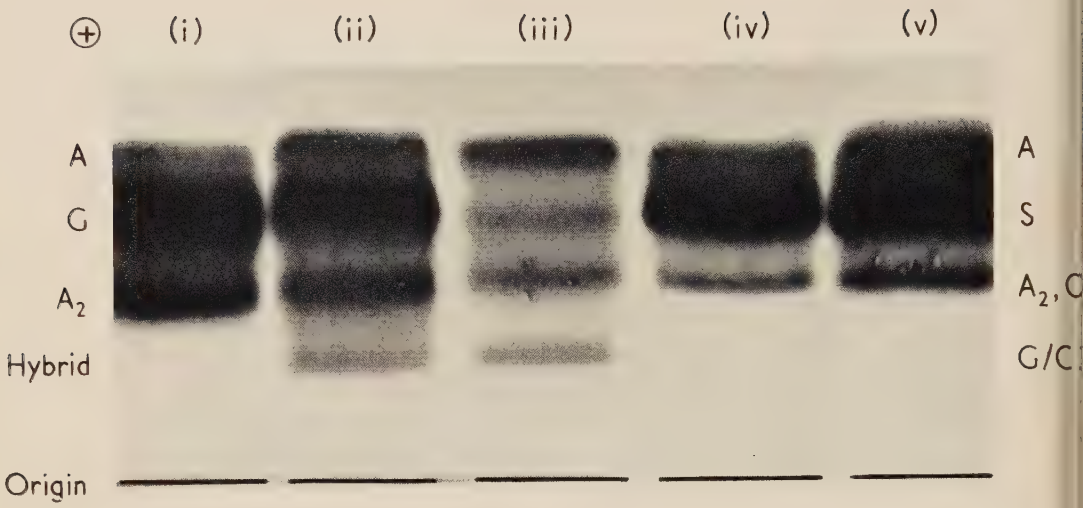


PLATE II. Hybridisation of haemoglobins A_2 and G_{Br}. Starch gel analysis in the discontinuous system. Migration time 2 hr, *o*-dianisidine stain. (i) Untreated mixture of Hb-A_2 and Hb-G_{Br}. (ii) Mixture of Hb-A_2 and Hb-G_{Br} dissociated for 96 hr and then recombined. (iii) Hb-A, G_{Br}, C and G_{Br}/C marker. (iv) Mixture of Hb-A_2 and Hb-S dissociated for 96 hr and then recombined. (v) Untreated mixture of Hb-A_2 and Hb-S.

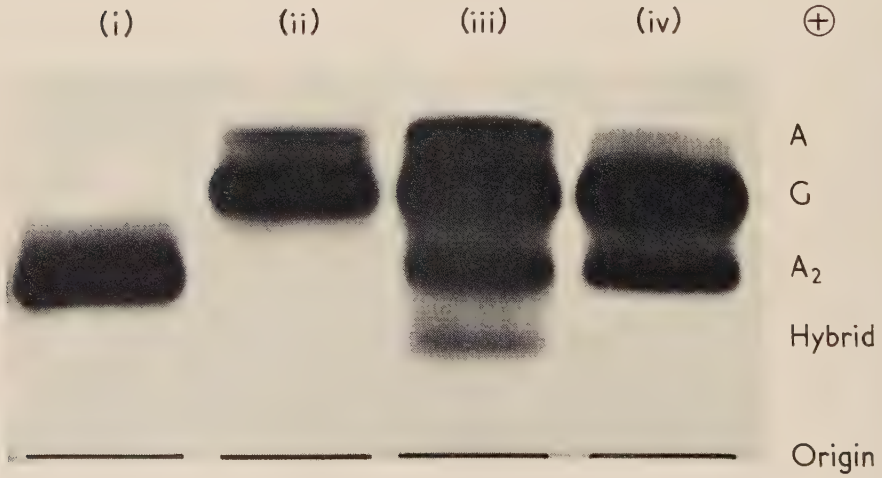


PLATE III. Hybridisation of haemoglobins A_2 and G_{Br}. Starch gel analysis in the discontinuous system. Migration time 2 hr, *o*-dianisidine stain. (i) Hb-A_2, (ii) Hb-G_{Br} and (iii) Hb-A_2 + Hb-G_{Br} all dissociated for 96 hr and then recombined. (iv) Untreated mixture of Hb-A_2 and Hb-G_{Br}.

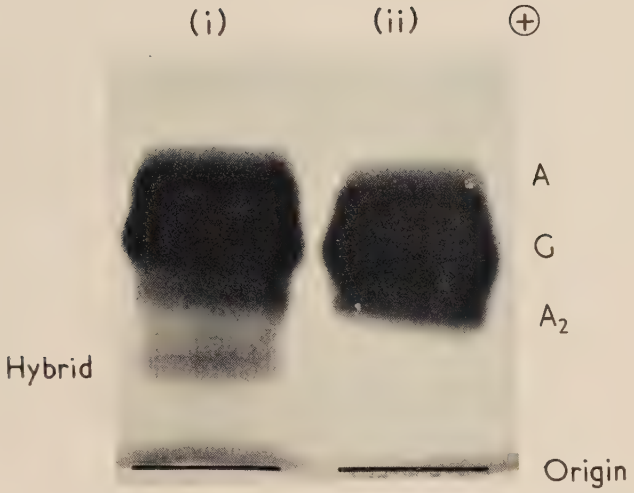


PLATE IV. Hybridisation of haemoglobins A_2 (from normal adult) and G_{Br}. Starch gel analysis in the discontinuous system. Migration time 2 hr, *o*-dianisidine stain. (i) Mixture of Hb-A_2 and Hb-G_{Br} dissociated and recombined. (ii) Untreated mixture of Hb-A_2 and Hb-G_{Br}.

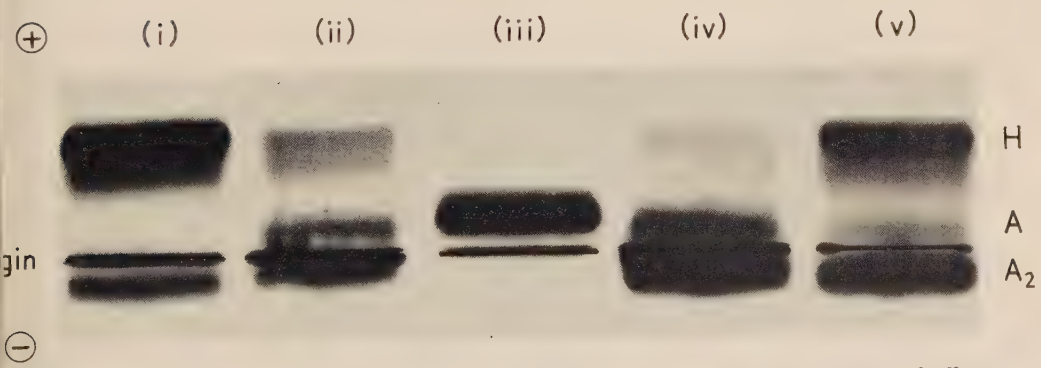


PLATE V. Hybridisation of haemoglobins A_2 and H. Starch gel analysis in phosphate buffer, gel pH 7·4. Migration time 2 hr, *o*-dianisidine stain. (i) Untreated mixture of Hb-A_2 and Hb-H in proportion of 1 : 5. (ii) Same mixture dissociated for 72 hr and then recombined. (iii) Hb-A. (iv) Mixture of Hb-A_2 and Hb-H in proportion of 1 : 1 dissociated for 72 hr and then recombined. (v) Same mixture untreated.

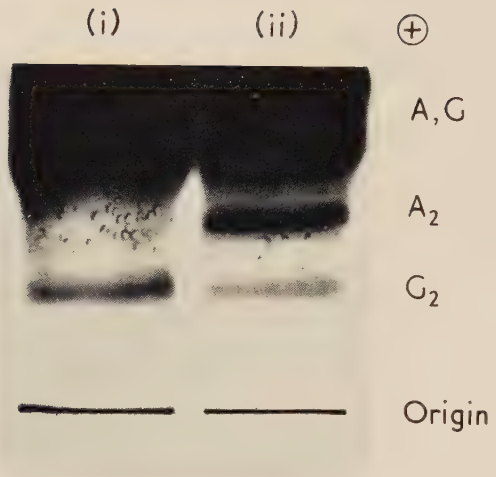


PLATE VI. Comparison of naturally occurring Hb-G_2 with the hybrid Hb-G/A_2 produced *in vitro*. Starch gel analysis in the discontinuous system. Migration time 2 hr, benzidine stain. (i) Natural Hb-G_2. (ii) *In vitro* hybrid Hb-G/A_2, same as Plate II(ii) with added Hb-A and Hb-G_{Br}.

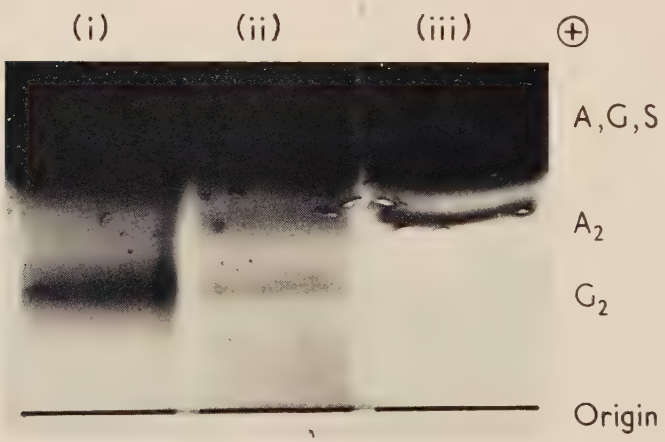


PLATE VII. Analysis in starch gel of (i) Hb-G_{Ib} disease showing Hb-G_2, (ii) Hb-G_2 trait showing Hb-A_2 and Hb-G_2 and (iii) Hb-S trait showing only Hb-A_2. Discontinuous system. Migration time 2 hr, *o*-dianisidine stain. Gel overloaded to emphasise minor components. Samples shown in (i) and (ii) have been stored for a prolonged time.

Hb-A_2 has, therefore, the composition $\alpha_2^A\delta_2^{A_2}$. The failure to form hybrid haemoglobins in the experiment where a mixture of Hb-A_2 and Hb-S (Plate II) was dissociated and recombined is in agreement with this formula.

This suggested molecular composition for Hb-A_2 is in agreement with the report by Ingram & Stretton (1960) that one of the differences between Hb-A_2 and Hb-A is in tryptic peptide 26 which is known to arise from the β-polypeptide chain. Moreover, Stretton & Ingram (personal communication) have recently shown that the tryptic peptides of the α-chains from Hb-A_2 are identical with those from the α-chains of Hb-A.

From a recent study of two individuals who carry the four haemoglobins A, G_{Br}, C and G_{Br}/C it was suggested (Raper *et al.*, 1960) that each type of subunit found in the haemoglobins of these individuals, namely α_2^A, $\alpha_2^{G_{Br}}$, β_2^A and β_2^C, is synthesized independently and that the final stage in haemoglobin synthesis is the combination of α_2 and β_2 subunits. It is probable that the synthesis of Hb-A_2 proceeds by the same mechanism, the $\delta_2^{A_2}$ subunits being synthesized on their own "template" before pairing up with α_2 subunits. If the individual is synthesizing only α_2^A subunits, then only the normal minor haemoglobin, Hb-A_2 is formed. However, if he is producing only abnormal α_2^G subunits, a new minor haemoglobin $\alpha_2^G\delta_2^{A_2}$ would replace Hb-A_2. This appears to be so, for the minor haemoglobin, Hb-G_2, found in the individual with Hb-G disease (Shooter *et al.*, 1960) migrates at a similar rate to the doubly abnormal hybrid, $\alpha_2^{G_{Br}}\delta_2^{A_2}$, produced *in vitro* from Hb-A_2 and Hb-G_{Br} (Plate VI). Finally, if the individual is synthesizing both α_2^A and α_2^G subunits, then both Hb-A_2 and Hb-G_2 ($\alpha_2^G\delta_2^{A_2}$) would be formed and this has been found in two unrelated carriers of the Hb-G trait (Plate VII). It therefore follows that any individual carrying an α-chain variant of Hb-A would also carry the corresponding minor component, $\alpha_2^X\delta_2^{A_2}$.

There are, therefore, two groups of Hb-A_2 variants. In the first group are those minor haemoglobins, for example Hb-A'_2 and Hb-B_2, which can occur without any concomitant abnormality of Hb-A, whereas the minor haemoglobins of the second group only occur in association with α-chain variants of Hb-A. This suggests that the synthesis of Hb-A_2 is controlled by two pairs of genes at separate loci, one being the α locus and the other specific for Hb-A_2, the δ-locus. The genotype relevant to the formation of Hb-A_2 in a normal individual would, therefore, be α^A/α^A, $\delta^{A_2}/\delta^{A_2}$. Applying the idea that each individual gene finds expression in its own "template" which in turn determines (in effect) the synthesis of a particular type of subunit (Raper *et al.*, 1960), it follows that a normal individual would only synthesize α_2^A and $\delta_2^{A_2}$ subunits leading to the formation of Hb-A_2. The individuals carrying the Hb-G_α trait would have the genotype α^A/α^G, $\delta^{A_2}/\delta^{A_2}$ resulting in the formation of α_2^A, α_2^G and $\delta_2^{A_2}$ subunits and the subsequent production of Hb-A_2 and Hb-G_2. In Hb-G_{Ib} disease the genotype would be α^G/α^G, $\delta^{A_2}/\delta^{A_2}$ leading to the synthesis of only Hb-G_2 ($\alpha_2^G\delta_2^{A_2}$). Mutation of a δ gene gives the genotype α^A/α^A, $\delta^{A_2}/\delta^{X_2}$; $\alpha_2^{A_2}$, $\delta_2^{A_2}$ and $\delta_2^{X_2}$ subunits are formed and Hb-A_2 and a new minor haemoglobin, $\alpha_2^A\delta_2^{X_2}$, result. The minor haemoglobins Hb-A'_2 and Hb-B_2 are presumably of this type.

We wish to thank Dr. H. Lehmann for a sample of haemoglobin G_α trait and for helpful discussion, and Dr. T. A. J. Prankerd for allowing us access to the patients with Hb-H disease and thalassaemia. We are also grateful to Dr. E. M. Crook for his interest and advice.

T

REFERENCES

Aksoy, M., Lehmann, H. & Eng, L. L. (1957). *Lancet*, i, 792.
Bailey, I. S. & Prankerd, T. A. J. (1958). *Brit. J. Haemat.* **4**, 150.
Bingle, J. P., Huehns, E. R. & Prankerd, T. A. J. (1958). *Brit. Med. J.* **2**, 1389.
Ceppellini, R. (1959). In *Ciba Found. Symp. Biochemistry of Human Genetics*, ed. by G. E. W. Wolstenholme & C. M. O'Connor, p. 134. London: Churchill.
Ceppellini, R., Kunkel, H. G. & Dunn, L. C. (1957). *Haemoglobin studies; Report on Research Activities of the Institute for the Study of Human Variation.* New York: Columbia University.
Everall, P. H. & Wright, G. H. (1958). *J. med. Lab. technol.* **15**, 209.
Gammack, D. B., Huehns, E. R., Shooter, E. M. & Gerald, P. S. (1960). *J. Mol. Biol.* **2**, 372.
Gammack, D. B., Huehns, E. R., Lehmann, H. & Shooter, E. M. (1961). *Acta Genetica Stat. Med.* **11**, 1.
Huisman, T. H. J., Martis, E. A. & Dozy, A. (1958). *J. Lab. Clin. Med.* **52**, 312.
Ingram, V. M. & Stretton, A. (1960). *Fed. Proc.* **19**, 217.
Itano, H. A. & Robinson, E. (1959). *Nature*, **183**, 1799.
Jones, R. T., Schroeder, W. A. & Vinograd, J. R. (1959). *J. Amer. Chem. Soc.* **81**, 4749.
Kunkel, H. G. (1954). In *Methods of Biochemical Analysis*, ed. by D. Glick, p. 141. New York: Interscience.
Kunkel, H. G. & Wallenius, G. (1955). *Science*, **122**, 288.
Raper, A. B., Gammack, D. B., Huehns, E. R. & Shooter, E. M. (1960). *Brit. Med. J.* **2**, 1257.
Shooter, E. M., Skinner, E. R., Garlick, J. P. & Barnicot, N. A. (1960). *Brit. J. Haemat.* **6**, 140.
Singer, K., Chernoff, A. I. & Singer, L. (1951). *Blood*, **6**, 413.

J. Mol. Biol. (1961) **3**, 263–266

Electron Microscopy of Unbroken DNA Molecules

Michael Beer

The Thomas C. Jenkins Department of Biophysics, The Johns Hopkins University, Baltimore, Maryland, U.S.A.

(*Received 10 January 1961*)

A method is described for transferring DNA from its solution to supporting films for electron microscopy as unbroken straight single molecules. The membrane used is a copolymer of styrene and vinylpyridine which has weakly basic ion exchange properties. DNA is transferred to the membrane when it is streaked along the solution. The amount transferred depends on the pH, ionic strength and concentration of the DNA solution.

1. Introduction

Attempts to relate genetic maps to the molecular fine structure of nucleic acids as revealed in the electron microscope require methods for obtaining electron micrographs of whole nucleic acid molecules. In the past only fragments of these giant macromolecules have been observed in the electron microscope. In this paper a method is described for transferring DNA molecules from their solution to a supporting membrane so that the molecules remain largely unfragmented and are very straight and parallel to each other.

The supporting film used is a copolymer of styrene and 4-vinylpyridine similar to that described by Birbeck & Stacey (1959). The specimen grid coated with this membrane is streaked along the surface of the solution containing the DNA molecules. The membrane acts as an ion exchange resin and the attachment of the molecules can be controlled by adjusting the pH and the ionic strength of the DNA solution. In this way the attachment can be made as gentle as required. The very high degree of orientation and straightening of the molecules results from the flow gradient accompanying movement of the grid.

2. Materials and Methods

(a) *Preparation of copolymer*

Styrene monomer was obtained from Eastman Organic Chemicals, Department of Distillation Products. 4-vinylpyridine was obtained from Reilly Tar and Chemical Corporation, New York. The chemicals were vacuum distilled and then distilled in the proportions desired into a glass cylinder containing enough benzoyl peroxide to make it 0·1%. It was preferable to distil the styrene first since the vinylpyridine is denser and better mixing occurs if it is condensed into the lighter styrene. The container was sealed under vacuum and polymerization allowed to proceed at 70°C for 12 hr. At the end of this time the polymer was hard. One per cent solutions were made in chloroform and the polymer was purified by re-precipitation as described by Fuoss & Cathers (1948).

(b) *Preparation of films*

Films were made from 0·1% solution of copolymer in chloroform. Clean glass slides were dipped in the solution and drained vertically on absorbent paper until dry. The films

were stripped soon after they were made and were picked up on a grid either from above or below. Those picked up from below were consistently smooth whereas those scooped from above sometimes were rough. Films could be stripped from mica with difficulty but no significant improvement in smoothness was obtained at the surface formed against the mica.

(c) *Transfer of DNA to grids*

DNA was prepared from T4 phage by the method of Thomas & Berns (1961). Solutions of nucleic acid in 0·1 M-phosphate buffer were dialysed for 2 hr against the desired buffers on a special dialysing apparatus. In it 0·2 ml. of the solution to be dialysed rests in a thin layer on a membrane which is continuously washed from below by fresh dialysing solution. The dialysing membrane was purchased from the Visking Company, Chicago, Illinois, and was boiled 5 min in 1% sodium bicarbonate, then cooled and rinsed in deionized water.

The DNA was transferred to the membrane on the grid by streaking the grid, membrane side down, along the surface of the DNA solution horizontally for a distance of 2 cm at a speed of about 2 cm/sec. The direction of streaking was noted on the grid. The excess liquid was immediately drawn off by placing the edge of the grid to absorbent tissue so that the flow direction of the receding excess liquid was the same as the direction of streaking. If this precaution was not taken some of the fibrils on the grid lay parallel to the streaking direction and others to the direction of the flow of the withdrawn excess liquid. Clearly there was a significant number of DNA molecules left in the excess liquid.

(d) *Examination*

Grids were shadowed with platinum at an angle of 5 to 1 in a direction perpendicular to the streaking direction and were then lightly coated with carbon. Serial photographs of the same molecule were most conveniently taken by orienting the grids in such a manner that the direction of the fibrils on the viewing screen was parallel to one of its sides.

3. Results

The density of fibrils picked up by the ion exchange membrane depended on the concentration of DNA, the pH of the suspending solution, the ionic strength of the suspending solution and the concentration of the vinylpyridine in the copolymer.

(a) *Recovery of DNA from volatile buffers*

The volatile buffer systems used were ammonium acetate-acetic acid or ammonium acetate-ammonia. Unless stated otherwise the solutions of DNA had an optical density of 0·3 at 260 mμ.

At buffer concentrations of 0·1 M and at pH 5·8 a very large amount of DNA was attached to the membrane and it was impossible to follow individual molecules. The orientation was not uniform and some of the fibrils were apparently aggregates of several DNA molecules. A typical portion of a grid is shown in Plate I.

At volatile buffer concentrations of 0·1 M and pH 6·1 few molecules were attached. They were all very accurately parallel, rather straight and could be followed for many microns. Their diameter as deduced from the dimension of the shadow was about 20 Å, indicating that they were single molecules. Plates II to VIII show serial photographs of the same strand. Arrows indicate corresponding points in the successive pictures. The length of the molecule is about 35 microns.

At volatile buffer concentrations of 0·1 M and pH 6·4 or higher only a negligible amount of DNA was attached to the membrane.

Using the same volatile buffers but at concentrations of 0·01 M with the same concentration of DNA it was found that DNA was attached to the membrane more extensively than at the higher ionic strengths. Also the pH optimum was less sharp and acceptable amounts of DNA were attached to the grid in the pH range 7·5 to 8.

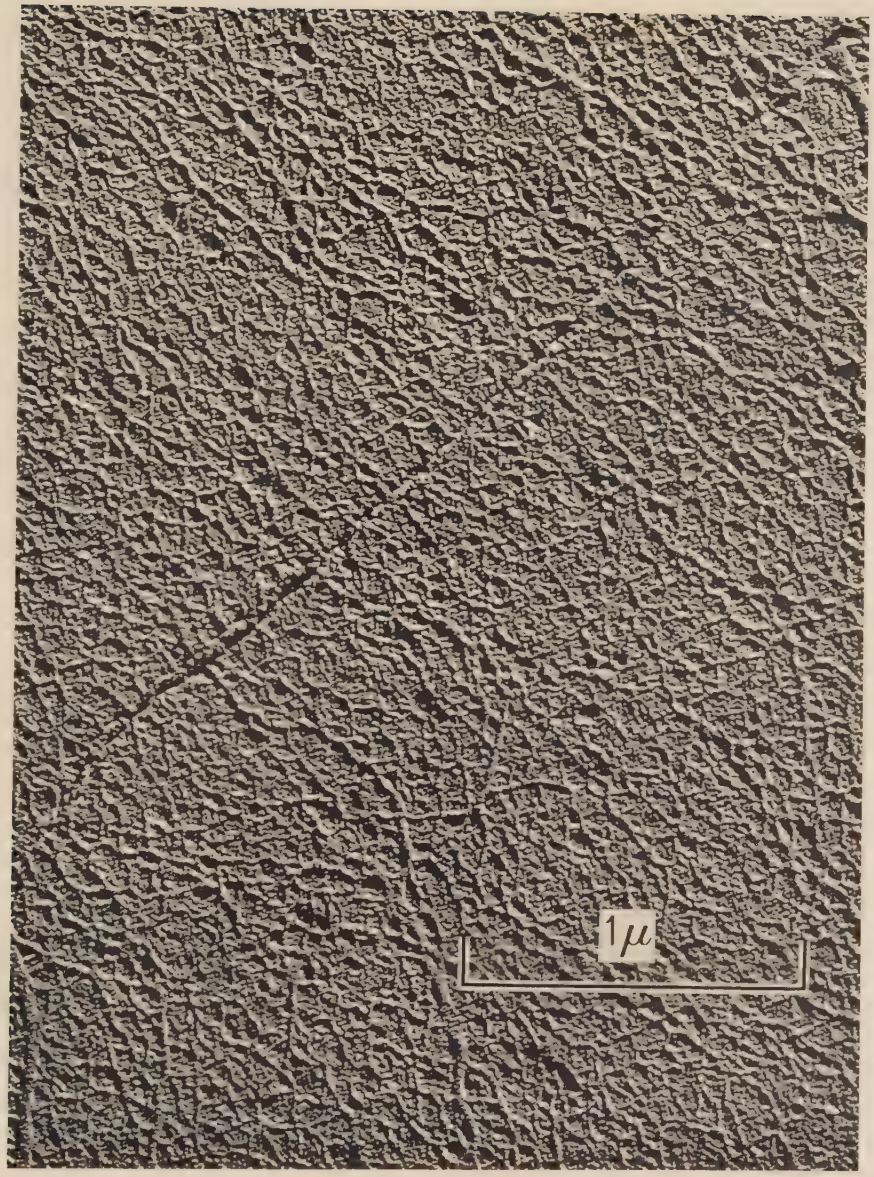


PLATE I. Appearance of specimen when pH of DNA solution was 5·8. The density of DNA is very high and the fibres are irregular. ×45.000.

[*To face page* 264

PLATE II

PLATES II to VIII. Adjacent overlapping micrographs of the same DNA molecule. Corresponding points of successive pictures are circled and are pointed out by arrows for easy recognition of continuity. × 58,000.

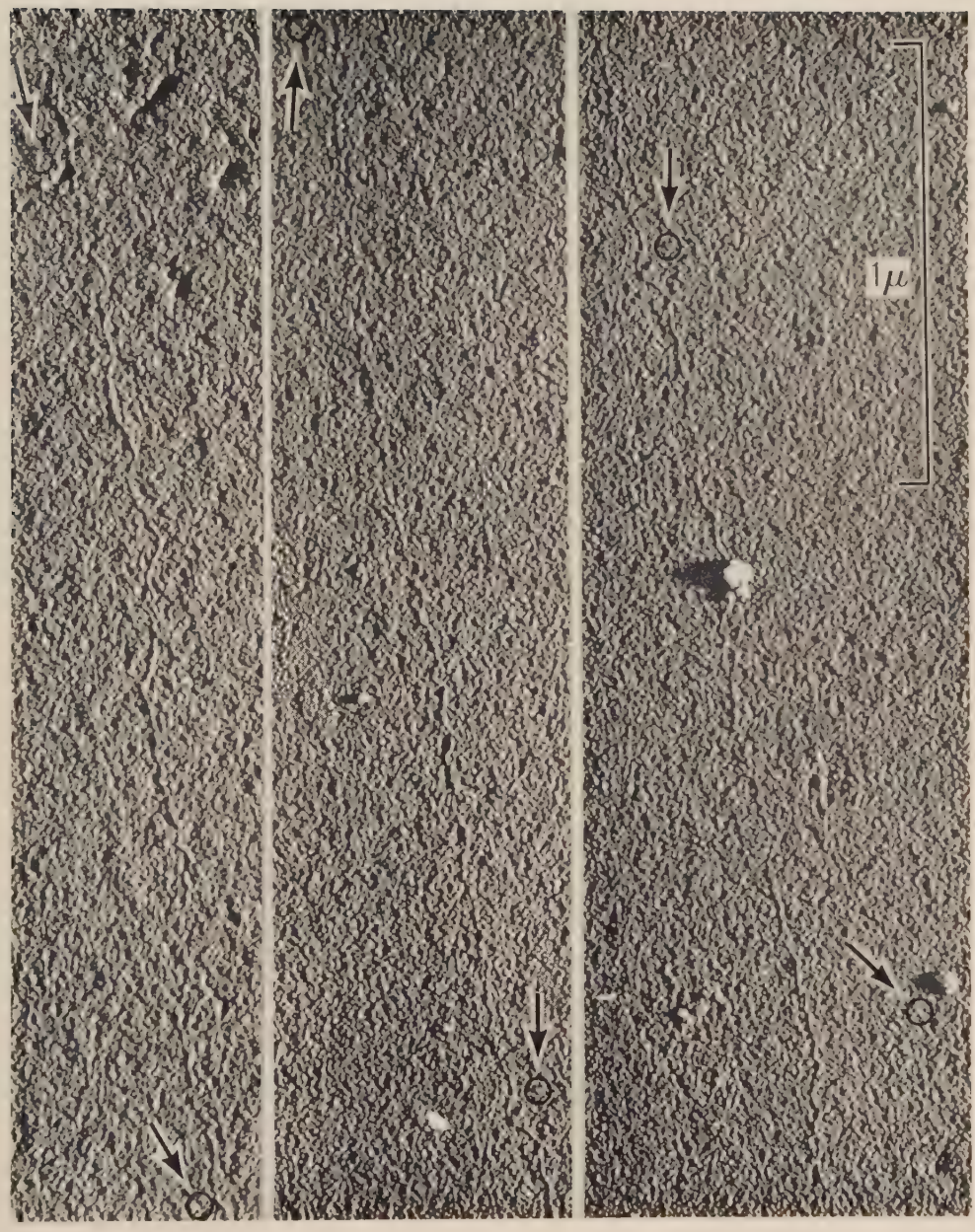


PLATE III

PLATE IV

Plate V

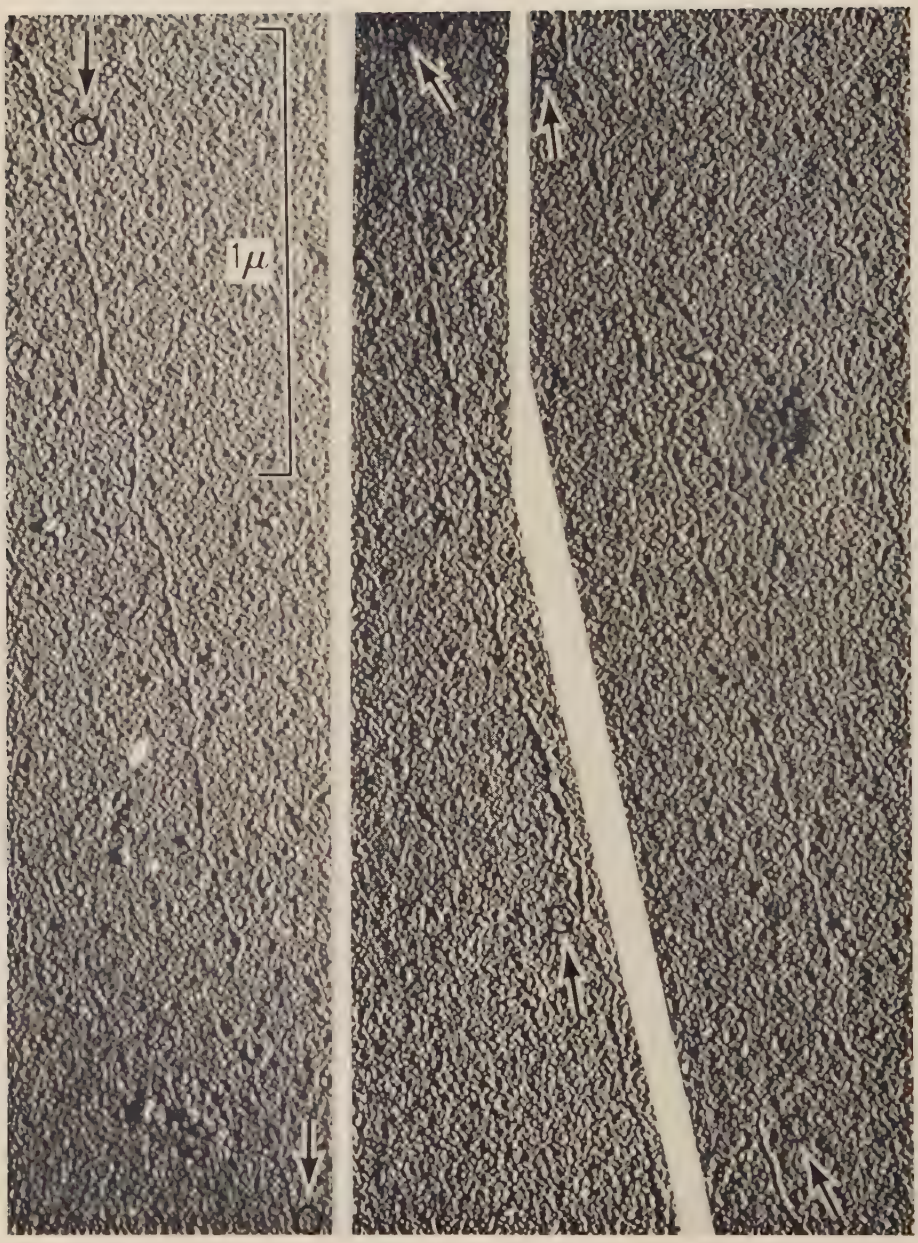


PLATE VI

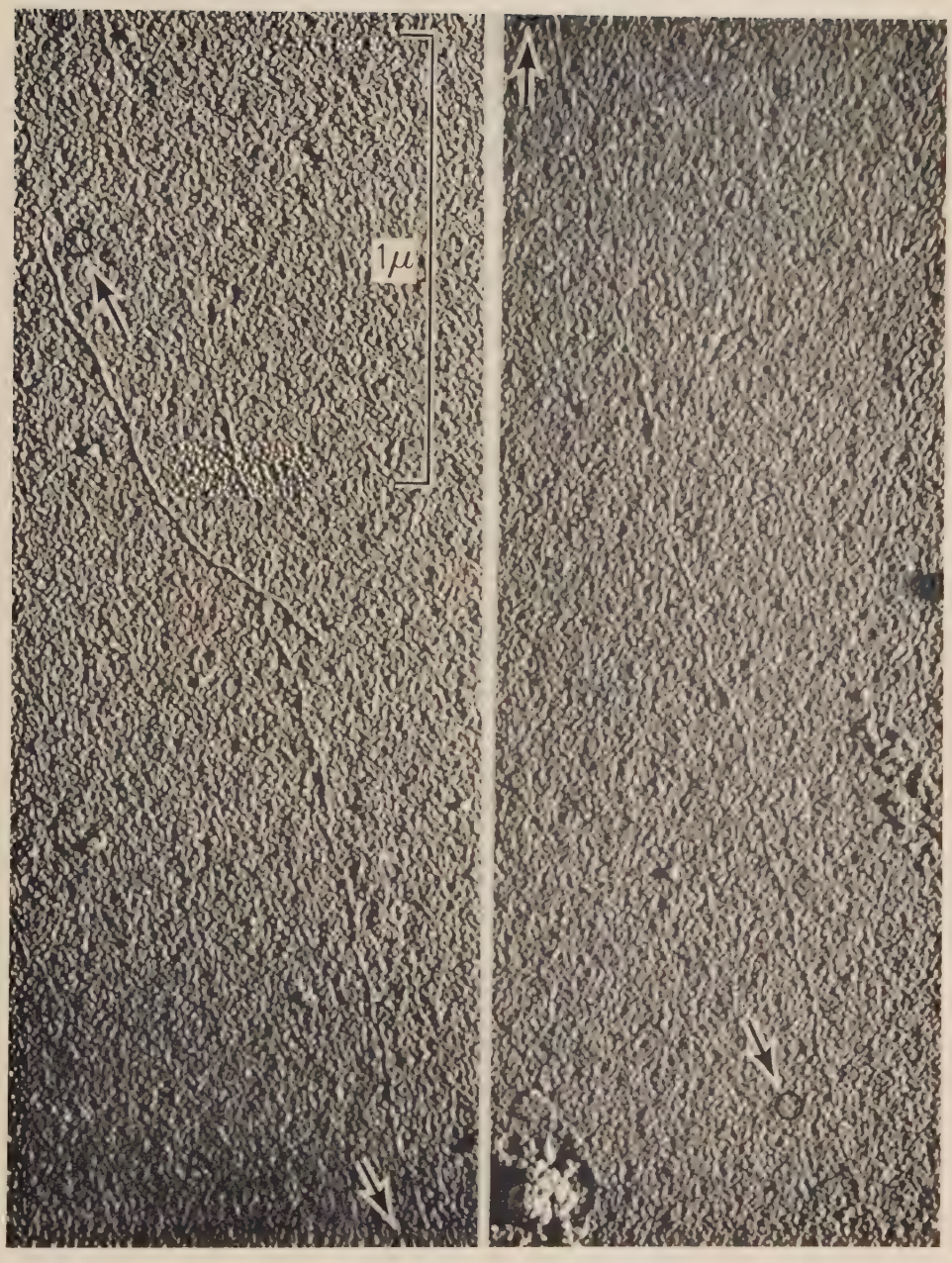


Plate VII

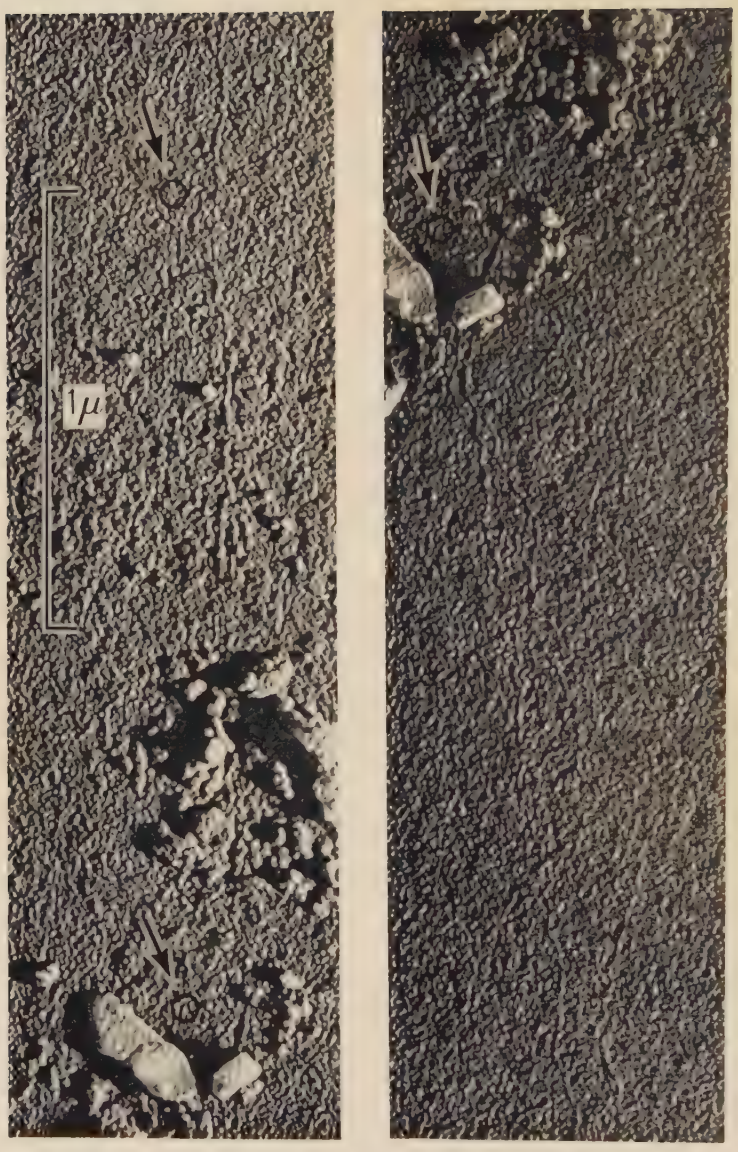


PLATE VIII

(b) *Recovery of DNA molecules from non-volatile buffer*

If DNA is transferred to the membranes from a non-volatile buffer solution the grids must be washed to remove the salts before examination. Using 0·1 M-phosphate buffer it was found that the amount of DNA left after washing was similar to that obtained from the volatile buffer at the same pH. However, at the optimum pH the molecules were disoriented. At lower pH's the orientation was better but the density of molecules too high. A satisfactory distribution of well-oriented DNA was obtained by reducing the pH of the buffer solution to 5·5 and reducing the concentration of the DNA in the non-volatile buffer tenfold, so that the optical density at 260 mμ was 0·03. The washing procedure consisted of streaking the wet grid, immediately after loading with DNA, along the surface of water pH adjusted to 5·0 with acetic acid.

(c) *Effect of concentration of vinylpyridine in copolymer*

Copolymer films containing, 10, 25, and 40% vinylpyridine have been made. The amount of DNA attached to the grid was very roughly proportional to the abundance of the vinylpyridine.

4. Discussion

The phenomena described in this paper can be understood when it is recognized that DNA is a negatively charged poly-electrolyte and the styrene-vinylpyridine copolymer has weakly basic ion exchange characteristics. The pK of the phosphate groups of DNA is about 2. Therefore the phosphates are essentially fully ionized and carry one negative charge each at pH above 5. The pK of pyridine is about 5·2. Assuming that the pK is not greatly altered by incorporation into the copolymer, at pH 6·2 one-twentieth of the pyridine groups are positively charged; the rest are uncharged. As the pH is raised the positive charge on the film is rapidly reduced. The decrease in the attachment of DNA as the pH of its solution was raised is a manifestation of the decreasing positive charge on the film.

When the buffer concentration was lowered from 0·1 M to 0·01 M, more DNA was attached to the film. This increase must be attributed to the greater attractive force between the poly-electrolyte and the membrane when the concentration of "gegen"-ions is low.

In the experiments with non-volatile buffer the amount of DNA attached to the film could be kept the same by lowering simultaneously both the pH and the concentration of DNA in the solution. The fibrils attached at lower pH values were held more tightly and did not become disoriented during washing.

Finally there remains the question of artifact: are the long strands single molecules? They are certainly obtained even when their separation on the grids is many microns. Therefore their length is not simply the result of accidental overlap during deposition. It is conceivable that fibres tend to be deposited along ones already attached to the film thus giving longitudinal aggregates rather than single molecules. Long overlaps can be ruled out by the apparent constancy of the diameter; however, short overlaps up to perhaps 0·1 micron could escape detection. Though this possibility is less likely, it cannot at present be rigorously ruled out and must await a statistical study of the lengths of fibres.

The author wishes to acknowledge the generous gift by Dr. C. A. Thomas, Jr. of some very high molecular weight T4 DNA; also the apparatus for rapid dialysis of small samples was designed and kindly lent by Dr. Thomas. Once again the able technical assistance of Miss Doris Naimark must be acknowledged. Dr. D. W. Robinson's help with the vacuum distillation was much appreciated. This research was financed by grants from the National Science Foundation and The National Institutes of Health.

REFERENCES

Birbeck, M. S. C. & Stacey, K. A. (1959). *J. Biophys. Biochem. Cytol.* **5**, 167.
Fuoss, R. M. & Cathers, C. I. (1948). *J. Polymer. Sci.* **4**, 97.
Thomas, C. A., Jr. & Berns, K. I. (1961). *J. Mol. Biol.* **3**, 277.

J. Mol. Biol. (1961) **3**, 267–276

Phage DNA Subunits in the Phage Precursor Pool †

A. W. Kozinski‡ and Hisao Uchida§

Virus Laboratory, University of California, Berkeley, U.S.A.

(*Received 25 January 1961*)

In the presence of chloramphenicol, phage precursor nucleic acid (DNA) can be allowed to accumulate to form a large intrabacterial pool, while protein synthesis has been temporarily arrested. Radiophosphorus, ^{32}P, of high specific activity was introduced at various stages and, by following the sensitivity to the radioactive ^{32}P decay of the final progeny population obtained after removal of the antibiotic, the origin of DNA phosphorus of individual phage particles which were derived from the DNA pool has been investigated. The experiments showed that radioactive decay inactivated the great majority of progeny phages, and that DNA synthesized at different points in time appeared together in the same phage individuals. The findings suggest either that the phage DNA is composed of *subunits* which are assembled into the entire viral DNA complement only upon maturation of the phages, or that, if the pool already contains the entire phage DNA complement synthesized in advance of its maturation into protein membranes, *exchange* of homologous subunits of the DNA with each other may occur during their sojourn in the pool.

1. Introduction

At the outset of intracellular growth, T2 and T4 bacteriophage particles inject their DNA into the bacterial host cell (Hershey & Chase, 1952). The parental phage DNA is then replicated several hundredfold and, after the termination of the eclipse period, replica phage DNA units are encapsulated into replica phage protein units to constitute the structurally intact, infective offspring viruses of the progeny generation. Three experimentally distinguishable sources contribute to the phosphorus atoms found in the DNA of the progeny viruses: phosphorylated constituents of the host bacterium, particularly the host DNA, present in the cells prior to their infection; phosphorus compounds assimilated from the growth medium after infection; and the DNA of the parent virus. The manner of distribution of these phosphorus atoms over populations of intact progeny phage maturing at different times indicates that, prior to maturation, the phage precursor DNA units exist in an intrabacterial *pool*, from which they are removed more or less at random by the maturation process (Stent & Maaløe, 1953; Hershey, 1953). The purpose of the present work was to investigate whether the entire DNA complement of a single phage particle is synthesized and maintained as an integral structure in this pool, or whether the pool is composed of DNA molecules, which are assembled to form the DNA complement of a particular phage only upon maturation of the progeny virus.

† Supported by a grant from the National Cancer Institute of the National Institutes of Health, United States Public Health Service.

‡ Present address: Department of Biophysics, Johns Hopkins University, Baltimore, Maryland, U.S.A.

§ Present address: Max Planck Institut für Biologie, Tübingen, Germany.

Experiments by Tomizawa (1958) have already shown that DNA molecules contained in any single phage particle may have been synthesized at different times of the eclipse, i.e. that the phage precursor DNA destined for a given virus does not seem to consist of a unitary structure constructed during a short time interval. This inference was based on Tomizawa's demonstration that, upon u.v. irradiation of phage-infected bacteria at late stages of the eclipse period, a considerable proportion of the phosphorus atoms assimilated from the medium only *after* irradiation ultimately find their way into progeny viruses carrying u.v. lesions in their DNA. More recent experiments with T4 phage (Koziniski, 1961) have shown that parental to progeny transfer of phage DNA is accompanied by the fragmentation of parental DNA molecules. The experiments reported here lead to a similar conclusion, in that they show that radioactive decay inactivates the great majority of progeny particles issuing from infected bacteria to which ^{32}P of high specific activity has been added at late stages of intercellular phage growth. This is to say, even though much of the phage precursor DNA has already been synthesized in non-labeled form prior to the addition of the radioisotope to the culture, very few of the virus particles ultimately matured draw only non-radioactive molecules for their DNA complement. Thus it seems that the phage precursor DNA is not a single structure which maintains its molecular integrity during the sojourn in the intrabacterial pool.

2. Materials and Methods

Bacteriophages. T2(H)r^+ and T2(H)r_7 (Hershey & Rotman, 1948) were used in the present study.

Bacteria. Escherichia coli, B/r (Witkin, 1951) was the host strain of the growth experiments. *E. coli*, BB (McFall & Stent, 1958) was the indicator strain in the phage assays, carried out on tryptone agar plates (Adams, 1950).

Medium. The tris-glucose medium described by Hershey (1955) was modified by omitting KH_2PO_4 and adding 5 mg/ml. Difco casamino acids. The medium thus modified contains 5·0 μg P per ml. The same medium without glucose or casamino acids was used as phage adsorption buffer, and will be referred to as tris-buffer. Casamino acids-glycerol medium (Fraser & Jerrel, 1953) was used for the storage of ^{32}P-labeled phage stocks.

Radiochemical techniques. The procedures of Stent & Fuerst (1955) were followed for preparing highly ^{32}P-labeled media, except that the modified tris-glucose medium was employed in place of H medium. Carrier-free radiophosphorus as $H_3{}^{32}PO_4$ was supplied by the Radiochemical Centre, Amersham, England.

Chloramphenicol. Crystalline chloramphenicol was obtained from Parke, Davis and Co.

Temperature. Unless stated otherwise, elapsed time periods in this paper refer to incubation at 37°C.

3. Experimental

(a) *Distribution of the bacterial phosphorus contribution over the progeny phage*

Approximately 30% of the phosphorus of the DNA of the phage progeny is derived from phosphorylated components of the bacterial host cell, assimilated prior to infection. The following experiment was carried out to study how this bacterial phosphorus contribution is distributed over the phage progeny under the conditions of growth employed in experiments to be reported below. For this purpose, a culture of non-labeled bacteria was infected with T2 phage and ^{32}P of high specific activity

added to the growth medium immediately after infection. Nine minutes later, chloramphenicol was added to the medium and the culture incubated further for some 40 min. During this time, synthesis of phage precursor DNA proceeds in the bacteria, while the inhibition of protein synthesis by the antibiotic prevents the formation of mature phage progeny (Hershey & Melechen, 1957; Tomizawa, 1958). At the 50th minute after infection, by which time a considerable pool of phage precursor DNA had accumulated in the infected cells, chloramphenicol was removed and incubation of the culture in the radioactive growth medium continued for various times, until an ever increasing number of infective progeny phages had made their intracellular appearance. After their release by artificial lysis of the bacteria (Fredericq, 1952), the ^{32}P-labeled progeny populations were stored at 4°C and assayed for their viability after various amounts of radioactive decay. A parallel control experiment was also carried out, in which labeled, instead of non-labeled, bacteria were infected and incubated until the release of mature progeny phage.

The results of this experiment are shown in Fig. 1, where curve A represents the survival of phage progeny grown on initially labeled bacteria. Curve B represents the survival of phage progeny grown on initially unlabeled bacteria and released by chloroform lysis, 85 min after the onset of phage growth, when an average of 5 mature progeny phages per infective center were present. It can be seen that radioactive decay inactivates both phage populations with simple exponential kinetics, the rate of inactivation of the phages released from the initially unlabeled bacteria being only 0·7 of the rate of inactivation of the phage released from the initially labeled bacteria. This difference in rates shows that the phages of curve B obtained about 30% of their phosphorus from the unlabeled phosphorylated constituents of the host bacteria, in good agreement with previous estimates of the magnitude of the bacterial phosphorus contribution (Cohen, 1948; Kozloff & Putnam, 1950; Labaw, 1951). Similar inactivation kinetics to those shown in curve B were also found for phage progeny released from initially unlabeled bacteria 100 and 135 min after infection, when 20 and 55 mature progeny per infective center were present.

The simple exponential kinetics of curve B show that the bacterial contribution is distributed very uniformly over the progeny population, so that less than 1% of the phage progeny can contain an appreciably greater than average bacterial contribution of unlabeled phosphorus. At first sight, these results seem to contradict earlier findings which indicated that the bacterial contribution appears preferentially in the first phages to mature after the end of the eclipse period (Weed & Cohen, 1951; Stent & Maaløe, 1953). However, these earlier experiments involved normal phage growth without exposure of the infected culture to the antibiotic. It must be inferred, therefore, that very efficient mixing of phosphorus assimilated before and after infection takes place in the large phage DNA precursor pool which builds up in the presence of chloramphenicol.

(b) *Contribution of phosphorus assimilated after chloramphenicol removal*

In the experiment just described considerable synthesis of phage DNA occurs during the time elapsing between removal of the antibiotic and intracellular appearance of the mature phage progeny. An experiment was carried out to assess the importance of this post-chloramphenicol phosphorus contribution to the DNA of the phage progeny. For this purpose, non-labeled bacteria were infected with T2 phage and then

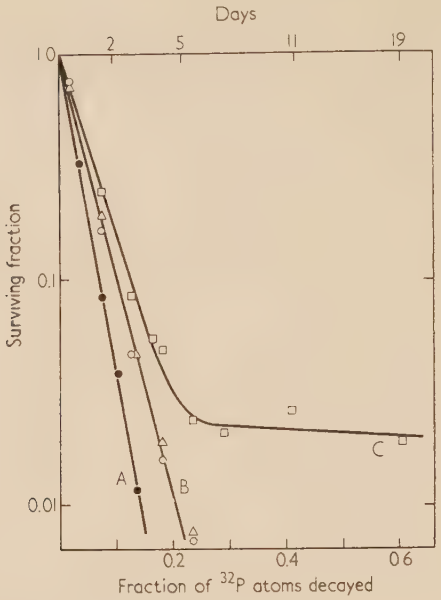


FIG. 1. Distribution of the bacterial phosphorus contribution over T2 progeny; survival with radioactive decay of phage grown under various labeling programs. *Curve A*: Host bacteria and medium ^{32}P-labeled throughout phage growth. *Curve B*: Host bacteria unlabeled and medium ^{32}P-labeled from infection until (a) lysis (○) and (b) removal of chloramphenicol (Δ). Chloroform-induced lysis at $t = 85$ min. *Curve C*: Same as part (b) of curve B, except chloroform-induced lysis at $t = 135$ min.

Procedure

Curve A. A culture of B/r bacteria was grown to a density of 5×10^7 bacteria/ml. in tris-glucose medium labeled at a specific ^{32}P activity of 290 mc/mg and infected at a multiplicity of 0·03 phages/cell. After incubation for 3 hr, the culture was diluted into cold casamino acids-glycerol medium, shaken with chloroform and the released progeny phage stored at 4°C and assayed from day to day.

Curves B and C. A culture of B/r bacteria was grown to a density of 5×10^8 bacteria/ml. in tris-glucose medium, centrifuged and resuspended in tris buffer to a density of 5×10^8 bacteria/ml. and starved by aeration for 30 min. This bacterial suspension was then infected with an average multiplicity of 4 T2r^+ phages/bacterium and phage adsorption allowed to take place for 5 min. The infected bacteria were separated from unadsorbed phage by two successive centrifugations and resuspensions of the pellets in tris buffer. Intracellular phage growth was initiated at $t = 0$ by adding a mixture of glucose and casamino acids to the infected suspension. At $t = 0{\cdot}5$ min, the culture was diluted 10-fold into tris-glucose medium labeled with ^{32}P giving a final specific activity of 290 mc/mg. Assay of a portion of the culture before and after treatment with chloroform showed that at this time less than one free parental phage per hundred infective centers was present. At $t = 9$ min chloramphenicol was added to the radioactive culture to a final concentration of 25 μg/ml. and incubation continued until $t = 50$ min, when an infectivity assay showed that 95% of the original infective centers had survived incubation in the antibiotic. At this time, one portion of the culture was diluted 1 : 200 into ^{32}P-labeled tris-glucose medium of the original radioactivity and another portion similarly diluted into non-labeled tris-glucose medium. Both portions were incubated further and samples from them diluted at $t = 85$ and $t = 135$ min into cold casamino acids-glycerol medium. These final dilutions were shaken with chloroform and the released progeny phages stored at 4°C and assayed from day to day. On the first day of the experiment, both portions contained an average of about 5 and 50 mature progeny phages per infective center at these two times of lysis.

treated exactly as in the preceding experiment except, that after removal of the antibiotic at the 50th minute of phage growth, the culture was incubated further in non-labeled, instead of in labeled, growth medium. Under these conditions, any phage

DNA synthesized after removal of the antibiotic will be non-labeled. The results of this experiment are also presented in Fig. 1, where it may be seen that the radioactive decay survival of phage progeny released 85 min after infection also follows curve B, exactly as that of the corresponding phages of the preceding experiment. Hence it may be inferred that very little DNA containing non-labeled phosphorus assimilated after removal of the antibiotic finds its way into the first 5 mature phage progeny formed. However, as is evident from curve C of Fig. 1, in this experiment the progeny population released 135 min after infection contains about 2% of individuals which are refractory to decay inactivation. Among the first 50 phages to mature, therefore, a few virus particles appear whose DNA contains mainly phosphorus atoms assimilated only after removal of the antibiotic.

(c) *Distribution over phage progeny of phosphorus assimilated after infection*

An experiment was carried out to study the distribution over the phage progeny of phosphorus atoms assimilated from the growth medium into the phage DNA precursor pool at different times *after* infection. For this purpose, non-labeled bacteria were infected with T2 phage and the infected culture treated exactly as in the preceding two experiments except that, instead of adding ^{32}P shortly after infection, radioisotope was introduced into the medium only at later stages of phage growth. That is, in the present experiment ^{32}P of high specific activity was added at the 15th and 30th min after infection to two different portions of a T2-infected culture in which protein synthesis had been arrested by chloramphenicol at the 9th min of phage growth. At the 50th min, the antibiotic, but not the ^{32}P, was again removed from the growth medium of both samples by dilution, the infective progeny phage released by chloroform lysis at the 85th min and the sensitivity of the progeny to ^{32}P decay examined. Under the conditions of this experiment, non-labeled phage DNA molecules arise in the phage precursor pool not only by utilization of the phosphorus contribution of the non-labeled bacterial host cells but also by post-infection assimilation of non-labeled phosphorus from the medium during the growth period preceding the addition of the label.

The results of this experiment are presented in Fig. 2 where curve B of Fig. 1 has been reproduced for comparison and curves D and E represent the survival of the progeny phage grown in cultures to which the label was added at the 15th and 30th min of growth, respectively. Curve D shows that radioactive decay inactivates at least 95% of the phage progeny whose intracellular development on non-labeled bacteria proceeded in the absence of the label for the first 15 min. The inactivation kinetics of this population is seen to follow a concave survival curve which, as a first approximation, may be considered to reflect the existence of two classes of progeny phage possessing different amounts of label in their DNA. One of these classes, constituting about 75% of the population, is inactivated at about 50% of the rate of death of the phages represented by curve B (to whose growth medium the label was added at the time of infection). The other class, constituting the remaining 25% of the population, is inactivated at about 10% the rate of death of the phages of curve B. Curve E shows that at least 80% of the phage progeny grown on non-labeled bacteria for 30 min before ^{32}P was added to the culture are sensitive to ^{32}P decay inactivation. In this case, the inactivation of about 90% of the decay sensitive individuals proceeds according to a simple exponential survival curve, whose slope is about 10% that of the rate of death of the population represented by curve B; the slight initial break in

the survival curve suggests that about 10% of the decay-sensitive phages are inactivated at a somewhat faster rate.

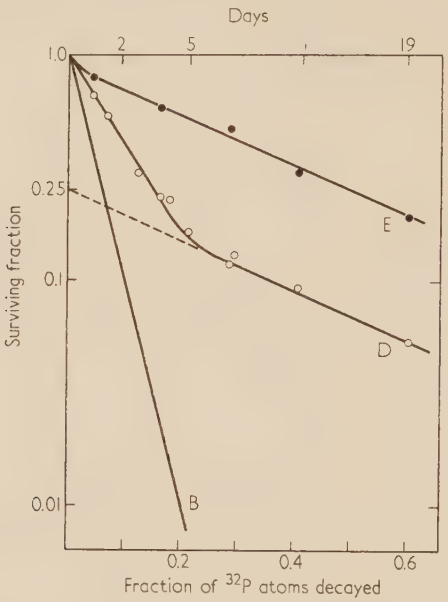


FIG. 2. Distribution among T2 progeny of phosphorus assimilated after infection; survival with radioactive decay of phage grown on unlabeled bacteria in culture medium to which ^{32}P was added at various times after infection. *Curve B*: label added at $t = 0{\cdot}5$ min (same as curve B of Fig. 1). *Curve D*: label added at $t = 15$ min. *Curve E*: label added at $t = 30$ min.

Procedure

B/r bacteria were grown, washed, starved, infected with T2r^+ phage and phage growth initiated after a 5 min adsorption period by addition of nutrient at $t = 0$, exactly as in the experiment of Fig. 1. At $t = 9$ min, chloramphenicol was added to the culture to a final concentration of 25 μg/ml. and incubation continued. At $t = 15$ min and $t = 30$ min, ^{32}P was added to a specific activity of 290 mc/mg to two different portions of this infected culture. At $t = 50$ min, both portions were diluted 1:200 into ^{32}P-labeled tris-glucose medium of the original specific activity and incubation continued until $t = 85$ min, when the cultures were diluted into cold casamino acids-glycerol medium and shaken with chloroform. The progeny phages released were stored at 4°C and assayed from day to day.

An exact quantitative appreciation of these results is difficult because there is no detailed knowledge of the metabolic pathway and pools through which the phage-precursor phosphorus is processed on its way from the inorganic phosphate of the medium to the phosphate ester bonds of the phage-precursor DNA. Nevertheless, it is possible to draw some semiquantitative conclusions. First of all, it is evident that the later ^{32}P is added to the infected culture, i.e. the more non-labeled phage precursor DNA that has already been synthesized in the phage-infected cells, the more slowly ^{32}P decay inactivates the phage progeny ultimately produced after removal of the antibiotic. Secondly, it is seen that even though an appreciable amount of non-labeled phage precursor DNA must have already been formed in the infected cell prior to addition of the label, very few progeny phage arise whose DNA complement is made up exclusively of non-labeled molecules, i.e. progeny phage that are entirely refractory to decay inactivation. The inference may thus be reached that the DNA complement of most of the first progeny phages to mature after removal of the

antibiotic is assembled from molecules synthesized at very different points in time during the building-up of the phage DNA precursor pool in the chloramphenicol-inhibited infected cells.

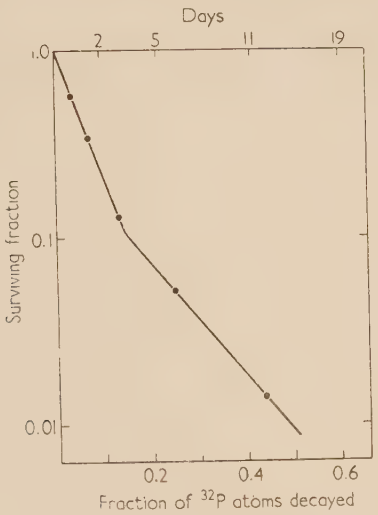


FIG. 3. Survival with radioactive decay of T2 phage grown on bacteria exposed to label while DNA synthesis was inhibited.

Procedure

B/r bacteria were grown, washed, starved, infected with T2r_7 phage, and phage growth initiated after a 5 min adsorption period by addition of nutrients at $t = 0$, exactly as in the experiment of Fig. 1. At $t = 0{\cdot}5$ min, chloramphenicol and ^{32}P were added to the infected culture to a final concentration of 250 μg/ml. and to a final specific activity of 350 mc/mg, respectively. At $t = 48$ min, enough $Na_2H^{31}PO_4$ was added to the culture to reduce the specific activity of the medium to 1 mc/mg. At $t = 50$ min, the culture was centrifuged, washed, resuspended in tris-glucose medium containing additional ^{31}P (0·5 mg/ml.) and returned to 37°C. Assays of portions of the culture before and after shaking with chloroform showed that at this point 60% of the original infective centers had survived centrifugation and exposure to the antibiotic, and that less than 1 free phage/100 infected centers was present. After incubation for 30 min more, the culture was diluted into cold casamino acids-glycerol medium and shaken with chloroform to liberate an average of about 2 progeny phages per infective center. The progeny phages were stored at 4°C and assayed from day to day.

(d) *Accumulation of phage precursor phosphorus in the absence of DNA synthesis*

Neither protein nor DNA is synthesized in T2-infected bacteria brought into contact with chloramphenicol immediately after infection (Tomizawa & Sunakawa, 1956; Hershey & Melechen, 1957). The following experiment was carried out to examine whether under these conditions the infective centers assimilate some phosphorus from the medium which can later be incorporated into the phage precursor DNA whose synthesis commences a few minutes after removal of the antibiotic and resumption of protein synthesis. For this purpose, ^{32}P and chloramphenicol were added to non-labeled bacteria shortly after their infection with T2 phage and the culture incubated for 50 min. At this time, ^{32}P and antibiotic were again removed and the culture incubated for 30 more minutes, when chloroform-induced lysis of the infected cells liberated an average of about 2 infective progeny phages per infective center. The sensitivity of these progeny to ^{32}P was then followed.

The result of this experiment is presented in Fig. 3 where it may be seen that radioactive decay rapidly inactivates the population of T2 progeny phages issuing from

bacteria exposed to the label only under conditions which do not permit any synthesis of DNA. The inactivation of the majority of this phage population proceeds at a rate of about 40% the rate of death of radioactive T2 stock uniformly ^{32}P-labeled at the specific activity of the labeled growth medium of this experiment. It may be concluded, therefore, that T2-infected bacteria in which early addition of chloramphenicol has prevented the onset of phage DNA synthesis do nevertheless assimilate phosphorus atoms from the growth medium that find their way into the DNA molecules of the progeny phage finally appearing after the removal of the antibiotic. Under such conditions, the newly assimilated phosphorus atoms must be present in the cell in some compounds other than DNA, prior to the onset of viral DNA synthesis.

It seems most likely that the provenance of these DNA precursor phosphorus atoms is the class of RNA molecules whose synthesis commences at the outset of phage growth in T2-infected bacteria, even when protein synthesis, and hence phage DNA synthesis, have been blocked by addition of chloramphenicol at the time of infection (Astrachan & Volkin, 1959). The presence of the antibiotic prevents the normal metabolic turnover of this phage-specific RNA, which, under these conditions accumulates within the infected cells. Upon removal of the antibiotic and resumption of protein synthesis, a rapid breakdown of the accumulated RNA molecules ensues. The result in Fig. 3 suggests that the phosphorylated products of the post-chloramphenicol breakdown of the phage-specific RNA serve as a source of phosphorus for the construction of phage precursor DNA, whose synthesis commences after removal of the antibiotic.

4. Discussion

The principal finding of the present experiments is that the DNA phosphorus of *individual* phage particles which obtain their germinal substance from the phage precursor pool built up while protein synthesis has been temporarily arrested is heterogeneous with respect to its origin. Some of the phosphorus, the bacterial contribution, is already present in the bacterial cells prior to infection, some of it is assimilated from the medium only much later, when considerable amounts of high-molecular weight DNA are already present in the phage precursor pool. Very little of the DNA complement of the first phages to mature after protein synthesis has been allowed to resume, however, contains phosphorus atoms assimilated after removal of the antibiotic. From this it must be concluded that in the phage precursor pool the DNA complement of vegetative T2 phages is not maintained as an integral stucture which, once synthesized, is destined to form the germinal substance of a particular progeny individual.

Two rather different mechanisms can be thought to prevent the preservation of the molecular identity of the phage precursor DNA. On the one hand, if synthesis of the T2 phage DNA in the pool proceeds according to the semiconservative replication mechanism of Watson & Crick, then the two complementary polynucleotide strands of every DNA molecule must, in any case, be of unequal "age," one strand always having been synthesized *de novo* at the last replication act and the other strand being the product of an earlier event (Delbrück & Stent, 1957). Hence in the experiment shown in Fig. 2 any non-labeled polynucleotide strands of DNA molecules synthesized before addition of the label would become the complementary partners of fully labeled polynucleotide strands in subsequent semi-conservative replication acts

occurring after the addition of the label, i.e. form part of DNA molecules labeled at half the maximum specific activity. In this way one could account for the appearance of that more highly labeled fraction of the phage population of curve D of Fig. 2 which is inactivated by ^{32}P decay at about half the maximum rate. On the other hand, semi-conservative replication of the phage DNA cannot readily explain the origin of the phage populations of curves D and E which are inactivated at only 10% the rate of death of the control population of curve B. Instead, one has to envisage that there is some other reason why DNA molecules synthesized at different points in time appear together in the same phage particle. The most likely explanation seems to be either that the phage precursor DNA pool is composed of *subunits* of the viral DNA assembled into the entire viral DNA complement only upon maturation of the virus (in a way formally similar to Luria's independent subunit hypothesis (Luria, 1947)), or that the pool already contains entire phage DNA complements synthesized as intact germinal structures which *exchange* homologous subunits with each other during their sojourn in the pool (as would be the case if breakage and reunion of preformed hereditary structures, and not "copy choice," were the basis of genetic recombination in phage). The data of Fig. 2 show that mixing of subunits in the precursor pool must be very efficient, since no more than a small minority of the first phage DNA complements to be withdrawn from the pool upon resumption of protein synthesis are comprised exclusively of DNA molecules synthesized before addition of the label. The complex nature of curves D and E in this figure reveals, nevertheless, that the mixing of the subunits is not so complete that the content of label of each particle is exactly equal to the average specific radioactivity of the pool. It is possible that at least one of these subunits is large.

These inferences are in good accord with earlier conclusions concerning the fate and distribution of the DNA transferred from parental to progeny phage. For it has been shown that the DNA complement of T2 and T4 phages is not transferred as an integral unit to the progeny but that it is instead dispersed over the DNA of several of the descendant viruses (Levinthal, 1956; Stent & Jerne, 1955; Hershey & Burgi, 1956; Kozinski, 1961). It seems likely that the processes producing the characteristic dispersal of the transferred parental DNA molecules in the vegetative pool are also responsible for the mixing of the newly synthesized phage DNA precursor molecules prior to the maturation into the infective progeny particles.

The authors are greatly indebted to Dr. G. S. Stent for invaluable criticism and advice, as well as for help in the preparation of the manuscript. Thanks are also due to Mrs. Margery Hoogs for technical assistance.

REFERENCES

Adams, M. H. (1950). *Meth. Med. Res.* **2**, 1.
Astrachan, L. & Volkin, E. (1959). *Biochim. biophys. Acta*, **32**, 449.
Cohen, S. S. (1948). *J. Biol. Chem.* **174**, 295.
Delbrück, M. & Stent, G. S. (1957). In *The Chemical Basis of Heredity*. Baltimore: The Johns Hopkins Press.
Fraser, D. & Jerrel, E. A. (1953). *J. Biol. Chem.* **205**, 291.
Fredericq, J. (1952). *C. R. Soc. Biol., Paris*, **146**, 327.
Hershey, A. D. (1953). *J. Gen. Physiol.* **37**, 1.
Hershey, A. D. (1955). *Virology*, **1**, 108.
Hershey, A. D. & Burgi, E. (1956). *Cold Spr. Harb. Symp. Quant. Biol.* **21**, 91.

Hershey, A. D. & Chase, M. (1952). *J. Gen. Physiol.* **36**, 39.
Hershey, A. D. & Melechen, N. E. (1957). *Virology*, **3**, 207.
Hershey, A. D. & Rotman, R. (1948). *Proc. Nat. Acad. Sci., Wash.* **34**, 89.
Kozinski, A. W. (1961). *Virology*, **13**, 124.
Kozloff, L. M. & Putnam, F. W. (1950). *J. Biol. Chem.* **182**, 229.
Labaw, L. W. (1951). *J. Bact.* **62**, 169.
Levinthal, C. (1956). *Proc. Nat. Acad. Sci., Wash.* **42**, 394.
Luria, S. E. (1947). *Proc. Nat. Acad. Sci., Wash.* **33**, 253.
McFall, E. & Stent, G. S. (1958). *J. Gen. Microbiol.* **18**, 346.
Stent, G. S. & Fuerst, C. R. (1955). *J. Gen. Physiol.* **38**, 441.
Stent, G. S. & Jerne, N. K. (1955). *Proc. Nat. Acad. Sci., Wash.* **41**, 704.
Stent, G. S. & Maaløe, O. (1953). *Biochim. biophys. Acta*, **10**, 55.
Tomizawa, J. (1958). *Virology*, **6**, 55.
Tomizawa, J. & Sunakawa, S. (1956). *J. Gen. Physiol.* **39**, 553.
Weed, L. L. & Cohen, S. S. (1951). *J. Biol. Chem.* **192**, 693.
Witkin, E. M. (1951). *Cold Spr. Harb. Symp. Quant. Biol.* **16**, 357.

J. Mol. Biol. (1961) **3**, 277–288

The Physical Characterization of DNA Molecules Released from T2 and T4 Bacteriophage

C. A. Thomas Jr. and K. I. Berns

The Biophysics Department, The Johns Hopkins University, Baltimore, Maryland, U.S.A.

(*Received 29 December 1960*)

The DNA molecules were released from T2 and T4 bacteriophage by a variety of procedures and examined by sedimentation equilibrium in CsCl density gradients and by sedimentation velocity and viscosity. The bands that resulted were quite Gaussian and very narrow, corresponding to molecular weights of about 62×10^6. These samples had sedimentation coefficients of about 57 to 63 s, and intrinsic viscosity values as high as 280 (100 cm^3/g). When these numbers are interpreted in the light of available theories, or empirical calibrations, they are in agreement with the molecular weight values derived from sedimentation equilibrium. The gradient equilibrium method itself has not yet been adequately calibrated in this molecular weight range, and there are certain important doubts concerning the experimental determination of the sedimentation coefficient and the intrinsic viscosity. If these doubts are set aside, it is possible to conclude that there are two molecules of DNA released from each phage particle, each of approximately the same molecular weight, 62×10^6.

Introduction

Interest in the organization of the DNA in T-even bacteriophage stems from two observations. First, it has been demonstrated (Hershey & Chase, 1952) that the infective agent is the DNA of the phage. Secondly the genetics of phage have been studied extensively (Adams, 1959). It is, therefore, of some importance to identify more precisely the molecular structure of the infective unit. Since these molecules contain the genetic information of the phage, a better understanding of their organization is essential to explain genetic results in physical and chemical terms.

In studying the organization of DNA in bacteriophage, one is faced with the choice of attempting to study the DNA in its folded condition *in situ*, or of extracting the DNA so that one may analyse the fragments in terms of molecular weight and composition. At present there appears to be no promising way of exploring the structure of the DNA inside the phage head, but the situation is not hopeless. Progress seems to be limited because of no adequate way of suitably orienting or crystallizing the tailed viruses.

If one chooses to disrupt the phage particle to release the DNA and protein fragments, the classical question arises of whether the disrupting procedure fractures the DNA molecules or establishes unnatural links. In addition, there are considerable practical difficulties in dealing with giant molecules liberated from phage. It was first recognised by Davison (1959) and now by a number of workers, including ourselves, that long DNA molecules are extremely sensitive to shear degradation (Mandell & Hershey, 1960; Thomas & Matheson, 1960). In an effort to circumvent these difficulties we have treated purified phage in a number of different ways to

release the DNA and compared the physical properties in solution in order to establish that the results do *not* depend on the mode of preparation. In these experiments, the solutions were transferred by pouring, or by mechanical devices using pipettes of large orifice in order to avoid shear degradation.

Another difficulty stems from the available methods for characterization of phage DNA. Most of the techniques such as light scattering, sedimentation velocity, intrinsic viscosity etc., either suffer from defects of principle or practice, or they have not yet been calibrated directly, or checked in the molecular weight range of 30×10^6 to 300×10^6. In this work we have concentrated on sedimentation equilibrium in cesium chloride density gradients and related these results to sedimentation velocity and intrinsic viscosity measurements on the same samples. The sedimentation equilibrium or "banding" procedure is relatively free of defects mentioned above, but the molecular weight values derived from it have not yet been adequately calibrated by other procedures; therefore, the absolute molecular weight derived from these experiments should be considered with some caution.

Radioautographic experiments on extracted phage DNA (Levinthal & Thomas, 1957; Thomas, 1959) had led to the idea of one large DNA molecule of 45×10^6 mol wt and five or six smaller ones of equal size for each phage particle. This view did not agree with the first gradient equilibrium experiments done by Meselson and coworkers (Meselson, Stahl & Vinograd, 1957) who concluded that the phage contained a uniform collection of molecules of 14×10^6 molecular weight. Finally the experiments of Hershey (Hershey & Burgi, 1960) lead to the idea that a phage particle contains a homogeneous collection of molecules, all of equal molecular weight. The experiments presented here lead to the same conclusion, but the molecular weight appears to be much higher—62×10^6. †

2. Methods

(a) *Phage growth and purification*

The phage used in most of these experiments was T4D or T2H. High titer lysates were prepared using *E. coli* strain BB growing in a glycerol-casamino acid medium (3XD) as previously described (Thomas & Matheson, 1960).

(b) *Release of DNA*

(i) *Osmotic shock.* Nineteen ml. of water at room temperature were poured into 1·0 ml. solution containing 2×10^{12} phage and 125 mg NaCl. This solution was then adjusted (without pipetting) to 0·10 M-NaCl, 0·032 M-PO_4, pH 7·2.

(ii) *Cell walls.* DNA-free cell walls were prepared (Barrington & Kozloff, 1956) and suspended in 0·10 M-NaCl and mixed with phage in the presence of about 10 μg/ml. L-tryptophan, so that there were approximately 10 phage for each bacterial equivalent of cell walls. In 30 min about 90% of the phage were inactivated. This solution was added directly to CsCl for density centrifugation.

(iii) *CdCN-lysine.* Approximately 5×10^{11} phage were added to a total of 5·1 ml. of CdCN solution (Kozloff & Lute, 1957) containing 10 μg/ml. of L-tryptophan. After 1 hr at 37°C or at 45°C, this solution was added to 2 vol 0·30 M-lysine, pH 9·0. After incubation for 1 hr at 37°C over 90% of the phage were inactivated and more than 85% of the DNA was rendered non-sedimentable under conditions that would pellet the protein ghosts. This solution was added directly to the CsCl solution for density centrifugation.

†See note added in proof, at end of paper.

(iv) *Centrifugation in CsCl.* A small amount (10%) of the DNA was released from the phage by mixing with 8·6 M-CsCl and spinning in the Spinco model E ultracentrifuge at 27,690 rev/min.

(v) *Phenol.* Shaking phage solutions with phenol releases the DNA and removes the phage protein. This is the procedure of Gierer & Schramm as perfected by Hershey (Mandell & Hershey, 1960). An important point is to use aqueous ether for the final removal of phenol. Approximately 5 to 8 extractions are necessary to eliminate the phenol completely. It was found that no ^{32}P is lost from the aqueous layer during the extraction. The ^{35}S content is reduced to about 0·1% of the intact phage (Mandell & Hershey, 1960). Our preparations contain less than 4% protein by the Folin test (Lowry, Rosebrough, Farr & Randell, 1951). Part of this color may result from a trace of phenol which cannot be removed by the ether extractions.

(c) *Intrinsic viscosity*

Viscosity measurements were performed in a multigradient capillary type viscometer described previously (Thomas & Matheson, 1960).

(d) *Sedimentation velocity*

Sedimentation velocity measurements were performed in 30 mm Epon cells at 31,410 rev/min in the Spinco model E ultracentrifuge. The DNA was diluted to approximately 15 μg/ml. in 0·10 M-NaCl, 0·032 M-PO_4, pH 7·2. The loading port was drilled out to accommodate a number 18 needle, and the DNA solution was allowed to flow in slowly over a period of a minute or more.

(e) *Sedimentation equilibrium*

(i) *Preparation.* Maywood "optical grade" CsCl was dissolved in 0·1% tris buffer (2-amino-2-hydroxymethylpropane-1:3-diol) to make a 59·6% w/w solution. In certain experiments 1% formaldehyde was present and the tris was replaced by 0·01 M-PO_4. One-tenth or 1/20 volume of DNA solution (in 0·01 M-NaCl, 0·032 M-PO_4) was added, and then sufficient water was added to bring the final refractive index to between 1·3992 and 1·4000 at 25°C. In those CsCl solutions which contained 1% formaldehyde, the measured refractive index was 0·0012 units higher. The "banding solutions" were introduced slowly into a 12 mm Epon cell *via* a number 18 needle. The centrifuge was run at 25°C for 60 to 120 hr depending on the rev/min. Equilibrium was tested by verifying the constancy of the band width on pictures taken 12 to 24 hr apart.

(ii) *Photography and densitometry of the band.* The conventional Spinco model E ultraviolet optics were used. The source was a G. E. H 85-C3 bulb followed by a special iodine filter (McFarland, Anderson, Nasim & McDonald, 1958) and the conventional Spinco chlorine-bromine filters. These filters reduce the intensity of visible light to less than 0·01%. Kodak Commercial film was exposed for various periods up to 160 sec, in order to determine the characteristic response of the film. The densitometry of the films was performed with a Joyce-Loebl microdensitometer. The blackening, or optical density, of the film was proportional to the logarithm of the time of exposure over the range of 40 to 160 sec. If the DNA concentration in the band was 10 times too high, there was no blackening at the peak of the band, a fact which testifies to the spectral purity. The result of spectral impurity and chromatic aberration is to produce "plateau-like" bands. This effect masks polydispersity effects when present.

The film was developed at 68°F in the following way: 5 min Kodak D-11 (full strength); 20 sec Kodak indicator stop; 6 min Kodak acid-fix; rinse; 2 min Kodak hypo clearing agent; 5 min wash; dip in "photoflo" solution; dip in distilled water; dry in a dust-free place (Meselson, personal communication, 1959).

(f) *Calculations*

(i) *Density and density gradient.* It becomes necessary to know the mean density at any place in the cell as well as the density gradient. A re-calculation of the thermodynamic expression for the dependence of the density gradient on angular velocity, assuming only CsCl and water were in the cell, reveals that if we denote $\frac{d\rho}{dr} = k\omega^2 r$, the value of k will be 8·27, 8·42, and $8{\cdot}86 \times 10^{-10}$, when ρ is 1·65, 1·70, and 1·76 g/cm^3 respectively. For this work a value of $8{\cdot}40 \times 10^{-10}$ was selected. On integration this expression yields:

$$\rho = \frac{1}{2} k\omega^2 \left[r^2 - \frac{1}{2}(r_m^2 + r_b^2) \right] + \bar{\rho} \qquad (1)$$

where r_m and r_b are the distance in cm to the meniscus and the bottom of the cell respectively. The quantity $\bar{\rho}$ is the density of the solution after thorough mixing. It is determined by measuring the refractive index of the solution both before and after a run. The refractive index is simply related to the density by the following empirical expression:

$$\rho\ (25°\text{C}) = 10{\cdot}8601\ \eta_D^{25} - 13{\cdot}4974$$

(Meselson, personal communication, 1959).

(ii) *Distribution of DNA in the band.* If the spectral purity and linearity of film response are adequate (as they are in this case) then the optical density of the DNA results in an alteration in the blackening of the film according to the expression $B_0 - B = \gamma D$, where γ is the slope of the response curve. Since the pen deflection of the densitometer is proportional to film blackening, this expression becomes $P_0 - P = b\gamma D$ where b is the constant of proportionality. In order to determine $b\gamma$ one must know how the pen deflection varies with exposure. This then allows the absolute determination of the optical density of DNA at any point in the cell, provided there is some region which is totally free of absorbing material.

(iii) *Plotting the results.* The distribution of DNA in the bands is determined as above and a graph is plotted of $\log \frac{C}{C(\max)}$ *vs* $\frac{W^2}{8\beta}$, where W is the band width. The expression

$$M = \frac{RT\rho(r_0)}{[d\rho/dr]_{r_0}\omega^2 r_0 \sigma^2} \qquad (2)$$

relates band width and molecular weight (Meselson *et al.*, 1957). It is convenient for plotting purposes to place together all values except σ^2, the band variance, into a constant β. Thus $\sigma^2 = \beta/M$. For a uniform collection of molecules of molecular weight M, the following expression should describe the distribution in the band:

$$\frac{D}{D(\max)} = \exp\left[-\frac{W^2 M}{8\beta} \right] \qquad (3)$$

In reporting the molecular weights a correction has been made from CsDNA to NaDNA by multiplying by $\frac{363}{474}$, the ratio of the average nucleotide unit in each case.

3. Results

(a) *Osmotic shock*

In Fig. 1 the results of a number of experiments on osmotically shocked T2 and T4 phage are presented. In the accompanying Table 1, additional information about these preparations is tabulated. In this figure the straight line labelled 14×10^6 depicts Meselson's data on osmotically shocked T4r_{240}. The dashed curve of intermediate slope represents the expectation for a mixture of one large molecule of 45×10^6 and five smaller ones of 16×10^6. The points falling nearest to this curve (△) represent

a run typical of six individual shock preparations including T4r_{240}. These solutions were handled in the 'normal fashion', that is by pipetting and loading the cell slowly with a number 22 needle. It can be shown that this kind of manipulation does not break most of the large star-forming DNA molecules, because they can be introduced and removed from the centrifuge cell using the number 22 needle (Thomas, 1959; Thomas & Knight, 1959).

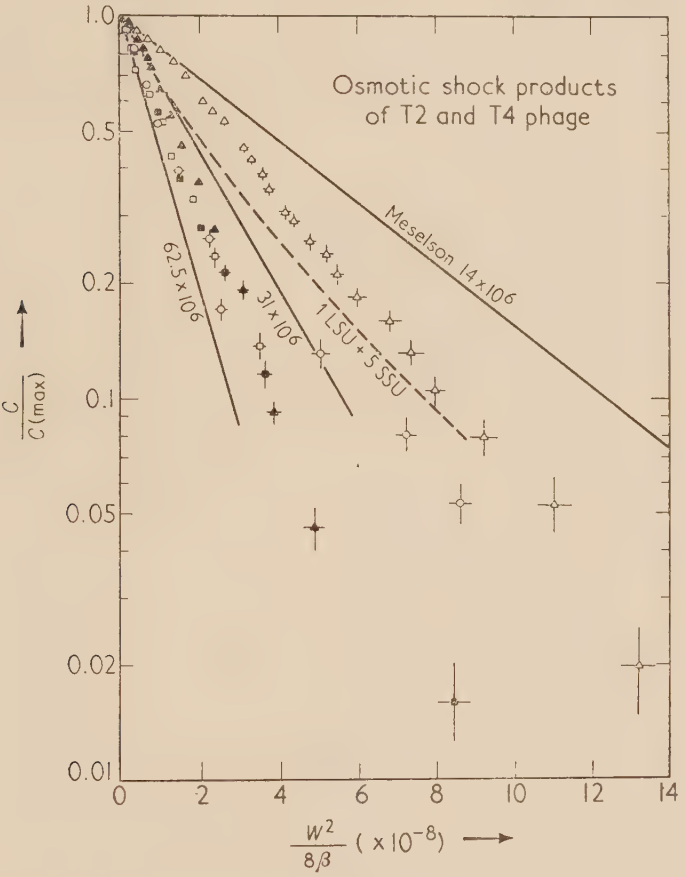


FIG. 1. Plot of band height *vs* (band width)$^2/8\beta$. Density centrifugation of phage DNA extracted by a variety of methods.

TABLE 1

Symbol	Run	Preparation	C(max) μg/ml.	$b\gamma$	ρ(g/cm^3)	Rev/min	$\beta\times10^{+4}$	$S\times10^{-13}$	$[\eta]$
△	64	osmotic shock	27·3	13·3	1·70	27,690	1·64	30	140–155
▲	172	osmotic shock	52·5	20·9	1·695	27,690	1·76	50	—
□	173	osmotic shock	14·0	18·1	1·697	27,690	1·60	57	—
■	83	cell wall	17·1	21·6	1·691	31,410	1·66	—	—
○	96	T4 + CsCl	8·0	21·9	1·694	27,690	1·69	—	—

On the other hand, if great care is taken with the transfer of the DNA solutions, namely by pouring or by using pipettes with very large orifices (*c.* 4 mm diameter), one gets even narrower bands and higher S values, for example runs number 83, 172 and 173.

In Fig. 1 are depicted two more bands; one resulting from the DNA released by the interaction of phage with bacterial cell walls (■), and the second resulting from the release of DNA from phage being centrifuged in a density gradient (○).

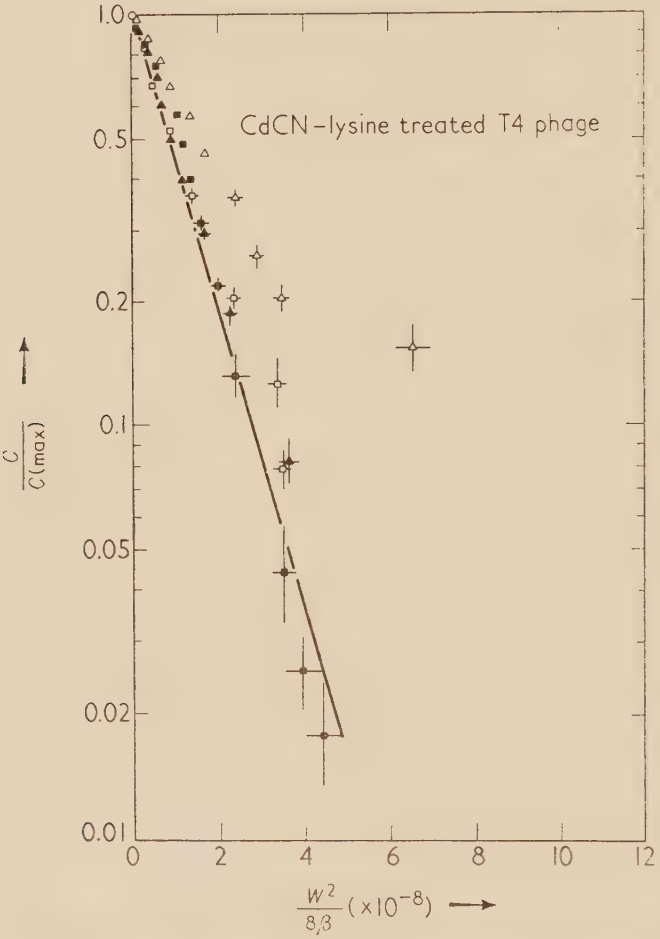


FIG. 2. Plot of band height *vs* (band width)$^2/8\beta$. Density centrifugation of T4 DNA extracted with CdCN-lysine.

(b) *CdCN—lysine*

Figure 2 and Table 2 summarize the results of disrupting phage with the procedure of Kozloff. This procedure probably allows the DNA to escape through the tail of the phage in a manner analogous to natural injection (Kozloff & Lute, 1957). The bands are somewhat narrower than the best osmotic shock products. An occasional preparation shows a broader distribution, perhaps resulting from inadvertent shear.

TABLE 2

Symbol	Run	C(max) μg/ml.	$b\gamma$	ρ(g/cm^3)	Rev/min	$\beta \times 10^{+4}$	[η]
▲	84	13·4	18·3	1·692	27,690	1·76	—
□	85	13·5	22·9	1·70	27,690	1·68	—
■	88	26·5	21·6	1·718†	27,690	1·50	—
△	89	21·2	23·5	1·678†	21,740	4·72	168–200

† The magnitude of the density in these experiments is suspect because of imperfect temperature control.

(c) *Phenol extraction*

Figure 3 and Table 3 summarize the results on phenol-extracted DNA. This procedure produces bands which are just as narrow as those resulting from the

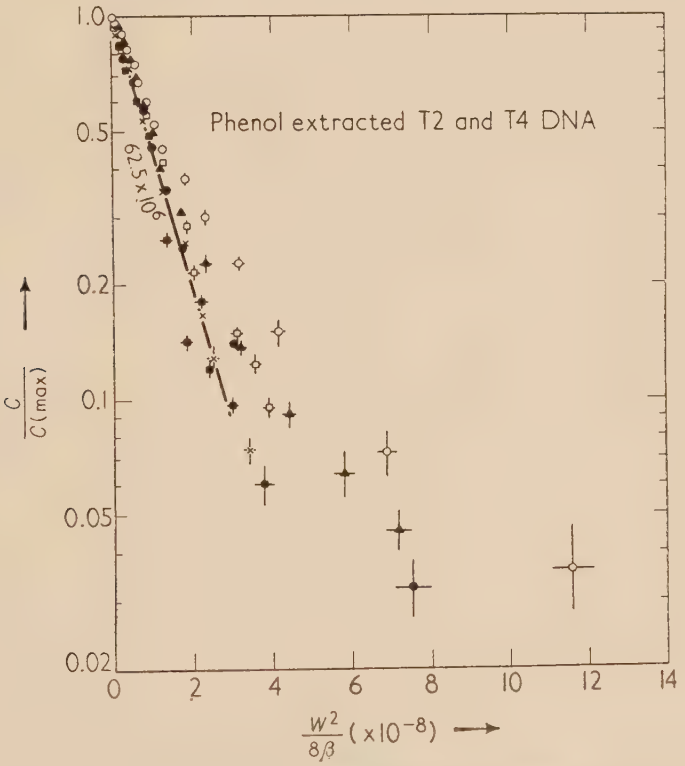


FIG. 3. Plot of band height *vs* (band width)^2/8β. Density centrifugation of T2 and T4 DNA extracted with phenol.

CdCN-lysine treatment. This treatment has the advantage that it not only releases the DNA but also removes the phage protein from the DNA solution. This graph includes a number of runs on DNA extracted by Hershey from T2H. This DNA shows an elutriation pattern typical of "whole" molecules with a negligible contamination of broken material (Mandell & Hershey, 1960).

Table 3

Symbol	Run	Sample	C(max) μg/ml.	$b\gamma$	ρ(g/cm^3)	Rev/min	$\beta\times10^{+4}$	$S\times10^{-13}$	$[\eta]$
○	219	T4D	26·0	21·2	1·692	31,410	1·19	—	—
▲	174	T4D	21·3	21·2	1·698	31,410	1·09	—	—
□	184	T2H	13·0	23·0	1·691	31,410	1·09	57·5	—
■	168	T2H	19·0	18·3	1·693	31,410	0·98	57·5	—
●	90	T2H	19·0	20·6	1·691	27,690	1·76	62·6	—
×	90	T2H	10·5	21·3	1·693	21,740	4·90	62·6	230–280

(d) *Stability*

The intrinsic viscosity of the CdCN-lysine treated phage, or of the phenol-extracted DNA, was not altered by heating to 70°C for 5 min nor by several days exposure to 80 μg/ml. of chymotrypsin and trypsin which had been demonstrated to be active against bovine serum albumin.

(e) *Determination of fraction of phosphorus in the band*

In order to determine what fraction of the DNA was truly in the band, and not floating or sinking out of sight, ^{32}P-labelled phage were osmotically disrupted, mixed with CsCl, and banded in the preparative swinging bucket rotor. Fractions were collected by piercing a hole in the "lusteroid" cup, and their radioactivity measured.

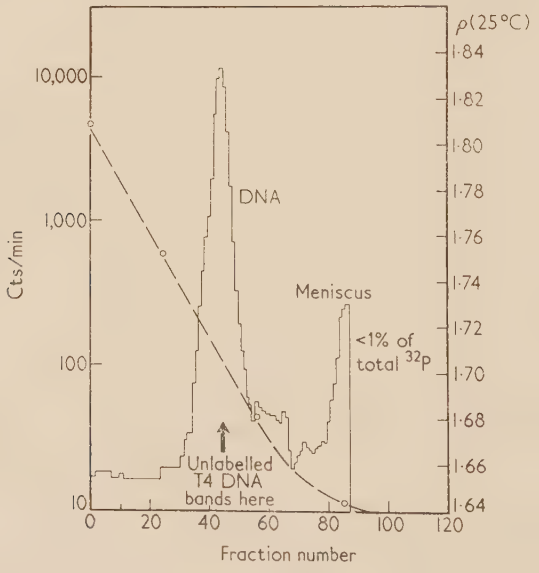


Fig. 4. Distribution of ^{32}P in density centrifugation run in Spinco model L of ^{32}P-labelled T4D DNA.

Less than 1% was found floating at the top. A typical experiment is shown in Fig. 4. One gets the same results with phenol-extracted DNA.

4. Discussion

(a) *Gradient equilibrium experiments*

The first question of genuine importance in banding experiments is whether or not all the DNA is in the band. The answer to this question can be obtained from the experiment depicted in Fig. 4. In independent experiments on the unseparated osmotic shock product, it was found that 98 to 99% of the ^{32}P banded between 1·68 and 1·725 g/cm^3, with the majority of the remainder floating at the meniscus. An equivalent experiment with the phenol-extracted DNA results in a fractionation profile with 99·9% of the ^{32}P banding between these densities. Much of this spread is a result of smearing during the fraction collecting process. Thus all but a small fraction of the ^{32}P can be demonstrated to be associated with macromolecules of mean density 1·693; thus the results from the analytical centrifuge runs concern the majority of the DNA, not just a fraction of it. Lehman has shown that at least 95% of this phosphorus is in the form of nucleotides (Lehman, 1960). This leaves a possible 5% of the phosphorus in some special macromolecular form.

One may estimate how much "protein" could be associated with some of the DNA molecules in the band. If we assume that the density of pure DNA is 1·693 (the average value of many runs on the phenol-extracted DNA), then these experiments indicate that there are less than 1% of the molecules having a density of less than 1·68. If we assume the density of protein to be 1·30, then this limit corresponds to 0·7% protein. Less than 1% of the DNA molecules contain as much as 1% of their weight in the buoyant protein. Repeating this kind of estimation on the analytical runs, one sees that 95% of the DNA bands with a density less than 0·002 g/cm^3 lighter than the mean value. This would mean that 5% of the molecules could be associated with as much as 0·08% protein of density 1·30. The value is certainly somewhat less than this, because such a contamination would tend to cause a slight asymmetry of the bands which is not observed. Moreover, these estimations assume no band spreading because of Brownian motion. Thus no more than 5% of the DNA molecules could be associated with buoyant protein of molecular weight 48,000, without escaping detection.

All of these remarks pertain to *heterogeneity* in protein content. It is still possible for each DNA molecule to have a unique fraction of protein which would change the density of *all* the molecules. If this is the case, the upper limit to this unique fraction is set by the contaminating ^{35}S of 0·1% (Mandell & Hershey, 1960) and the 4% protein by Folin reaction which is the maximum value displayed by any of our phenol-extracted preparations.

These arguments lead to the conclusion that the DNA consists of nucleotides over the majority of its length. The evidence that T-even phage DNA is double helical stems from two kinds of evidence. First its X-ray diffraction pattern conforms with DNAs of other origins which have received exhaustive crystallographic study (Hamilton *et al.*, 1959). Secondly, the characteristic 20 Å molecule can be seen and measured on mica grids (Thomas & Beer, unpublished experiments, 1958). These experiments mean that at least some of the DNA is duplex. To prove that the majority of the DNA is duplex, one looks to two other kinds of experiments. First, phage DNA has a characteristic, sharp melting point (Marmur & Doty, 1959) of 83°C with an associated increase (43%) in extinction coefficient (Berns & Thomas, 1960). Secondly, undenatured phage DNA does not react with formaldehyde at room temperature or at 45°C for 12 hr. Since the free nucleotides react rapidly at these temperatures showing a marked alteration in the extinction coefficient, these experiments put an

upper limit of 1 to 2% for the formaldehyde-reactable sites (amino groups) present in the native molecule. Finally, if there were a sizable fraction of single-chained material in the molecule, one would expect a shift in density (see next paper). Thus if the majority of the DNA were duplex, one could easily detect 10% of the molecules which have more than 4% of their nucleotides in the denser single-chained form. As with the discussion on contaminating protein, this does not rule out the possibility that each molecule has a single-chained region along 1 to 2% of its total length.

Figures 2, 3 and 4, along with their corresponding tables, suggest that the maximum molecular weight of the DNA released from T2 or T4 phage by any of the methods we have used is between 50 and 70×10^6. Lower molecular weights are considered to result from shear degradation. The data are thus in agreement with $62{\cdot}5 \times 10^6$ which would represent exactly half of the total DNA in the phage particle. A single molecule of $125 \pm 10 \times 10^6$, which is the best estimate of the total DNA in the phage (Cummings & Kozloff, 1960) is definitely not compatible with the data.†

(b) *Some unknown effects which influence the interpretation*

In this interpretation, we have assumed that there is no other contribution to the variance other than Brownian motion, whereas density heterogeneity would be expected to make a large contribution (Sueoka, 1959; Rolfe & Meselson, 1959; Marmur & Doty, 1959). The fact that these molecules come from a genetically pure lysate of phage and that there are only one or two molecules per phage particle, encourages the notion that they should be very similar in composition, and hence density (Sueoka, Marmur & Doty, 1959). In any event this effect would cause an underestimation of the molecular weight because density heterogeneity would cause the observed bands to be too broad, and this would be interpreted as a lower molecular weight. Another factor which could cause the bands to be of different shape, and the molecular weights derived from them to be in error, is preferential salt or water binding by the DNA. This "three component" effect has been, rightly or wrongly, ignored here.

A final word of caution should be added concerning the calibration of the banding technique. There are as yet no experiments published which directly relate the results of this method with the results of others; therefore, the question of absolute molecular weights must remain open for the moment. As an example of the kind of disturbing thing that one encounters, if these molecular weights are correct, and the DNA molecule is a linear duplex, then its contour length is about 33 microns. If the ratio of the contour length to the r.m.s. distance between the beginning and end points of the coil, R, is 5 : 1 (which is similar to that of lower molecular weight DNA (Eigner, 1960)), then the molecule occupies a spherical domain about six microns in radius or 12 microns in diameter. The band width (2σ) in most (but not all) of the experiments presented here is only about 260 microns, only 22 molecular diameters. At the peak, the band is only a few molecular diameters in width. Overlapping of the molecular domains is extensive. One wonders if the molecules can be considered independent kinetic units. The fact that DNA molecules *do* interact at these concentrations is clearly shown by the effect that "whole" molecules display in protecting against shear degradation (Hershey & Burgi, 1960). On the other hand, we can find no concentration effect on the band form; the derived molecular weight is not changed by an alteration in the concentration, or by changing the speed of the rotor.

† See note added in proof, at end of paper.

This is illustrated in Table 3, by runs 219, 169 and 90. If there is significant spectral impurity, this will result in an apparent alteration.

(c) *Polydispersity*

Finally, a word should be said about polydispersity. If a collection of molecules all have the same density, but differ in molecular weight, then the band should not be Gaussian, but should conform to the following expression:

$$\frac{C}{C_0} = \sum_i f_1 e^{-\frac{W^2 M_i}{8\beta}} \tag{4}$$

where W is the band width and C is the weight concentration of DNA molecules at the position corresponding to W. M_i is the molecular weight of the i^{th} molecular weight species, and f_i is a measure of their abundance.

$$f_i = \frac{w_i \sqrt{M_i}}{\sum_i w_i \sqrt{M_i}} \tag{5}$$

The quantity w_1 is the weight fraction of the i species.

Using these equations it is possible to estimate the amount of polydispersity. We assume only two components, for simplicity, one of which has a molecular weight of 62×10^6. Our "best" CdCN-lysine preparations could only have 29% by weight of 46×10^6 material, while the "worst" phenol preparations might have as much as 23% of 17×10^6. In all cases, there can be as much as 4% of the DNA with molecular weight as low as 5×10^6. None of these statements is compatible with the conclusions of the earlier radioautography which lead to the idea of large and small subunits.

(d) *Sedimentation velocity and viscosity*

The interpretation of sedimentation viscosity measurements on DNA solutions has been extensively investigated by Eigner (1960) who has measured S and $[\eta]$ over a 50-fold molecular weight range ending at 15×10^6. If one extrapolates his S-$[\eta]$ plots, one finds that an S of 55 corresponds to an intrinsic viscosity of 280. These are approximately the values obtained here, a fact which makes it seem likely that this phage DNA is similar to Eigner's bacterial DNAs of lower molecular weight. An expression derived by Eigner based partly on the Mandelkern-Flory equation (Mandelkern & Flory, 1952) and partly on an empirical calibration is shown below.

$$\log M_w = 4{\cdot}269 + 1{\cdot}33 \log (S_0 [\eta]^{1/3}) \tag{6}$$

On substituting in these values of S and $[\eta]$ one calculates a molecular weight of 52×10^6, a figure which is in reasonable agreement with that derived from the band widths. One may take this agreement as tending to support the molecular weight values derived from the banding experiments.

This support would have been stronger if it had not been for the appearance of a disturbing effect of rotor speed on the sedimentation coefficient of high molecular weight DNAs and dextrans (Ellias, 1959; Eigner, 1960; Hershey, unpublished experiments). Similar doubts have been cast upon the determination of the intrinsic viscosity of large molecules in capillary viscometers by the work of Schultz-Grunow (1958). If these effects turn out to be large, the molecular weights derived from S-$[\eta]$ calculations will have to be modified.

(e) *Conclusions*

Therefore, depending on the resolution of these experimental and conceptual problems, these molecular weights may be increased or decreased. At this point, the tentative conclusion is that each phage contains two DNA molecules, each of about 62×10^6 mol wt.† These results are not compatible with the earlier star counting on DNA molecules, and it is suspected that these were shear degraded fragments of larger DNA molecules.

We wish to thank Dr. A. D. Hershey for many helpful discussions and samples of his DNA. We are also grateful to Dr. Matthew Meselson, who has been very generous with much information as yet unpublished. It is a pleasure to acknowledge the support of the United States Atomic Energy Commission, Contract AT (30-1)-2119.

† *Note added in proof.*

Recent radioautographic determinations (Rubenstein, Thomas & Hershey, 1961) of the molecular weight of these unbroken DNA molecules has revealed that there is only *one* molecule per phage particle and that it comprises all of the DNA in the phage-120×10^6 mol. wt. units. Thus our fears concerning the calibration of the banding technique (and of sedimentation-viscosity) have proved justified: they are approximately a factor of two too low.

REFERENCES

Adams, M. H. (1959). *The Bacteriophages*, New York: Interscience.
Barrington, L. F. & Kozloff, L. M. (1956). *J. Biol. Chem.* **223**, 615.
Berns, K. I. & Thomas, C. A., Jr. (1961). *J. Mol. Biol.* **3**, 289.
Cummings, D. & Kozloff, L. M. (1960). *Biochim biophys. Acta*, **44**, 445.
Davison, P. F. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1560.
Eigner, J. (1960). Thesis, Harvard University.
Ellias, H. G. (1959). *Die Makromolekulare Chemie*, **33**, 166.
Hamilton, L. D., Barclay, R. K., Wilkins, M. H. F., Brown, G. L., Wilson, H. R., Marvin, D. A., Ephrussi-Taylor, H. & Simmons, N. S. (1959). *J. Biophys. Biochem. Cytol.* **5**, 397.
Hershey, A. D. & Burgi, E. (1960). *J. Mol. Biol.* **2**, 143.
Hershey, A. D. & Chase, M. (1952). *J. Gen. Physiol.* **36**, 39.
Kozloff, L. M. & Lute, M. (1957). *J. Biol. Chem.* **228**, 511.
Lehman, I. R. (1960). *J. Biol. Chem.* **235**, 1479.
Levinthal, C. & Thomas, C. A., Jr. (1957). *Biochim. biophys. Acta*, **23**, 453.
Lowry, C. H., Rosebrough, N. J., Farr, A. L. & Randell, R. J. (1951). *J. Biol. Chem.* **193**, 265.
McFarland, R. H., Anderson, R. A., Nasim, M. & McDonald, D. G. (1958). *Rev. Sci. Instrum.* **29**, 738.
Mandelkern, L. & Flory, P. J. (1952). *Chem. Phys.* **20**, 212.
Mandell, J. M. & Hershey, A. D. (1960). *Ann. Biochem.* **1**, 66.
Marmur, J. & Doty, P. (1959). *Nature*, **183**, 1427.
Meselson, M. S., Stahl, F. W. & Vinograd, J. (1957). *Proc. Nat. Acad. Sci., Wash.* **45**, 332.
Rolfe, R. & Meselson, M. S. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1039.
Rubenstein, I., Thomas, C.A., Jr. & Hershey, A. D. (1961). *Proc. Nat. Acad. Sci. Wash.* in the press.
Schultz-Grunow, F. (1958). *Rheologica Acta*, **1**, 289.
Sueoka, N. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1480.
Sueoka, N., Marmur, J. & Doty, P. (1959). *Nature*, **183**, 1429.
Thomas, C. A., Jr. (1959). *J. Gen. Physiol.* **42**, 503.
Thomas, C. A., Jr. & Knight, J. H. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 332.
Thomas, C. A., Jr. & Matheson, A. T. (1960). *J. Mol. Biol.* **2**, 5.

J. Mol. Biol. (1961) **3**, 289–300

A Study of Single Polynucleotide Chains Derived from T2 and T4 Bacteriophage

K. I. Berns and C. A. Thomas, Jr.

Department of Biophysics, The Johns Hopkins University, Baltimore, Maryland, U.S.A.

(*Received 29 December 1960, and in revised form 27 February 1961*)

The phenol-extracted DNA from phage was subjected to various procedures which disrupt the hydrogen bonds, but do not break any of the ester linkages. This denatured DNA reacts rapidly with formaldehyde and, once reacted, there is no tendency to "renature" or re-establish the former hydrogen bonded structure. These formaldehyde-reacted polynucleotide chains have been examined by density-gradient centrifugation and the band width and band form are compared with the same quantities measured on the undenatured molecules.

The results indicate that the molecular weight of the single polynucleotide chains is very high, and in no case lower than one-half that of the undenatured molecules from which they are derived. If this is correct, it means that the polynucleotide chains which make up the large undenatured molecules of 62×10^6 mol. wt. consist of only two continuously-linked polynucleotide chains. This conclusion bears on current ideas of genetic recombination.

1. Introduction

In the preceding paper we have come to the tentative conclusion that the DNA molecules released from T2 and T4 phage have a molecular weight of about 62×10^6, or half the DNA content of the entire phage (Thomas & Berns, 1961). It is now of some importance to examine the question of non-DNA, or unique polynucleotide linkages which might conceivably unite distinguishable macromolecular subunits into the very large structures mentioned above. The fact that the intrinsic viscosity is not changed by heating to submelting temperatures, or by prolonged reaction with proteases (Thomas & Berns, 1961), puts certain restrictions on the kind of linkages which might exist.

In this report we describe the results of a more profound treatment which was devised to test the hypothesis of single-chain breaks. If there existed interruptions in the polynucleotide chains making up the duplex, the molecule would be weaker at these special points. On exposure to mild mechanical shear, the molecule might be broken into discrete subunits, a possibility which could account for the origin of the large and small subunits found by radioautography (Thomas, 1959). There are other reasons for the interest in the possibility of these breaks. These have to do with current theories on the fragmentation and distribution of parental DNA to progeny phage particles during intracellular phage DNA synthesis (Stent, 1959, p. 265). If these interruptions occurred, they might bear on our understanding of the genetic recombination process (Levinthal, 1959).

The stoichiometric detection of a few interruptions or "breaks" per DNA molecule is not possible. On the other hand, a single break in a polynucleotide chain would cause a large change in molecular weight, provided, of course, that it was completely free from any others in solution. To obtain completely free polynucleotide chains

requires special conditions because one must be certain that all of the 100,000 hydrogen bonds between nucleotides in each molecule are in fact broken, and yet be sure that none of the ester linkages between nucleotides are broken by the treatment. Secondly, it is now known that following denaturation, there is a strong tendency for the hydrogen bonds to re-establish themselves, and for sizable regions of the duplex to be reconstituted (Doty, Marmur, Eigner & Schildkraut, 1960). This, or any non-specific aggregation, would lead to erroneous conclusions.

In this work we have denatured the DNA by a variety of methods and reacted the exposed amino groups with formaldehyde, measured the molecular weight of the formaldehyde-reacted polynucleotide chains, and then compared this molecular weight with that of the undenatured DNA in the same experiment. The rate of backbone breakage by these treatments was found to be small and negligible. The possibility of aggregation was made unlikely by introducing a few breaks with DNase and noting that this treatment did indeed cause the molecular weight of the polynucleotide chains to be reduced, although the molecular weight of the duplex was not greatly affected.

2. Methods

The methods and procedures employed here are essentially those described previously (Thomas & Berns, 1961). Phenol-extracted DNA was used throughout these experiments, with one exception.

Baker Analysed Reagent grade formaldehyde solution containing 37% CH_2O and 12·5% methanol was used without further purification. The methanol does not interfere with the reaction with purified nucleotides (Haselkorn, 1959).

Spectrophotometric measurements were made in the Cary spectrophotometer in a thermostated cell-jacket which could be maintained at any desired temperature up to 100°C.

The reduced specific viscosity (η_{sp}/c) was determined by a simple Ostwald-Fenske viscometer which had a flow time for water of about 50 sec.

All the sedimentation equilibrium procedures have been previously discussed (Thomas & Berns, 1961).

3. Results

(a) *The reaction of CH_2O with T4 DNA*

If one heats samples of T4 phage DNA in 0·10 M-NaCl, 0·032 M-PO_4, pH 7·2 to various temperatures and measures the optical density at 260 mμ at each temperature, one finds the characteristic increase in extinction coefficient (Marmur & Doty, 1959). On cooling the sample, even rapidly, a large fraction of this increase is recovered, and on re-heating one again finds the sharp increase at 83°C. In the presence of 1% CH_2O, however, there are two differences. The increase in extinction coefficient appears at a lower temperature, and a sample heated to any temperature displays very little or no recovery in the extinction coefficient. Therefore a series of samples was prepared containing 1% CH_2O, heated to various temperatures for 3 min, then chilled and the optical density remeasured. These results are shown in Fig. 1.

To explore this reaction further, the time course of the optical density change at 260 mμ was followed at a number of different temperatures. Approximately 85% of the observed density increase at this wavelength is due to thermal denaturation while 15% is contributed by the formaldehyde reaction with the bases. The results are shown in Fig. 2. The quantity

$$\log\left(1 - \frac{\text{O.D.}_t - \text{O.D.}_0}{\text{O.D.}_\infty - \text{O.D.}_0}\right)$$

should be proportional to the time for a first order reaction. We see that there is no detectable reaction at room temperature or at 45°C. At temperatures above 75°C

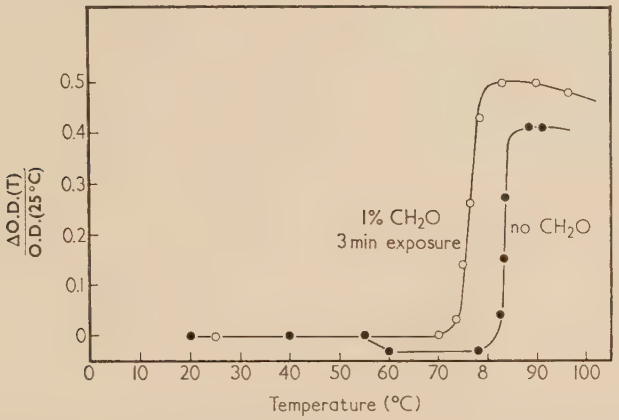


FIG. 1. Influence of CH_2O on hyperchromic effect. Comparison of the change in absorbance (260 mμ) as a function of temperature between T4 DNA in the presence of 1% CH_2O and in the absence of CH_2O. In both cases the solvent was 0·10 M-NaCl, 0·032 M-PO_4, pH 7·2.

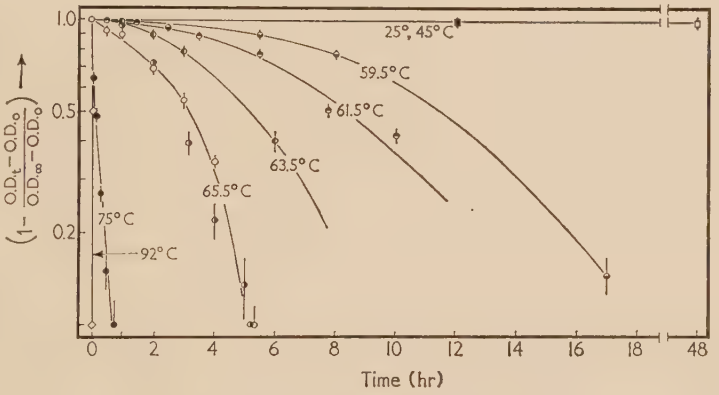


FIG. 2. The reaction of 1% CH_2O with T4 DNA at various temperatures. The relative change in absorbance (260 mμ) is plotted as a function of time in the presence of 1% CH_2O at different temperatures. The curve labelled 65·5°C represents two experiments. The first (●) was done at constant temperature in the Cary model 11 recording spectrophotometer and the other (○) was accomplished by removing a portion from a constant temperature water bath and quickly cooling and measuring the absorbance at 25°C. In all cases the solvent was 0·10 M-NaCl, 0·032 M-PO_4, pH 7·2.

the reaction is rapid and is apparently first order. At intermediate temperatures the rate of the reaction increases with time. This would imply that the number of sites made available for the reaction with formaldehyde is increasing with time. Leaving open for the moment the interpretation of this interesting result, the important conclusion for the subsequent experiments is that there is no detectable reaction below 45°C while there is a rapid reaction at 92°C.

The kinetics of denaturation in 15% CH_2O at 45°C also reveal this continuously increasing rate of denaturation. Thus the *rate* depends on time, temperature, and CH_2O concentration.

The reaction of CH_2O with DNA in 0·20 M-NaOH appears to be first order, with the reaction going 99% to equilibrium in about 2 min, when the formaldehyde is present to about 1%. As pointed out by Grossman (Grossman, Levine & Allison. 1961) the reaction appears to be base catalysed.

Some preliminary experiments on the reaction of 1% formaldehyde with the pure deoxynucleotides indicate that the kinetics are first order between 20°C and 95°C with activation energies of between — 15 to — 18 kcal/mole. In the presence of 1% CH_2O at 25°C the extent of the reaction at equilibrium is approximately 75 to 80%.

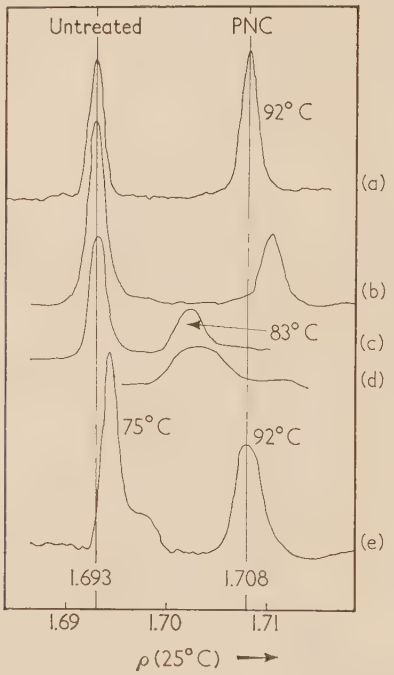


Fig. 3. Density gradient centrifugation of DNA preparations. The optical density of the DNA in the cell is plotted vertically, while the density (as obtained from the position in the cell) is plotted horizontally. PNC = Polynucleotide chains.

(a) Native DNA mixed with denatured DNA (92°C, 3 min, 1% CH_2O).
(b) Native DNA mixed with denatured DNA (pH 12·2, 30 min, 1% CH_2O).
(c) Native DNA mixed with "partially denatured" DNA (83°C, 3 min, 1% CH_2O).
(d) "Partially denatured" DNA (83°C, 3 min, 1% CH_2O).
(e) "Partially denatured" DNA (75°C, 3 min, 1% CH_2O) mixed with denatured DNA (92°C, 3 min, 1% CH_2O).

(b) *Banding of CH_2O-reacted polynucleotide chains*

When the formaldehyde-reacted DNA is mixed with an equal amount of unreacted DNA and then banded in a CsCl density gradient, one gets two sharp bands. A series of experiments is shown in Fig. 3. In the first experiments of this kind, formaldehyde was not included in the CsCl solutions and a broad smear was found between the density positions marked by the two bands in Fig. 3. It is likely that this is due to partial re-formation of the duplex molecules when the CH_2O is removed. Some

preliminary spectrophotometric experiments also support this view. Therefore in all subsequent experiments reported here, 1% CH_2O is present in all solutions.

Taking the density of the native material to be 1·693 g/cm^3, those samples heated to 92°C for 3 min are 0·015 units higher. The same is true if the polynucleotide chains are prepared by heating with 15% CH_2O for 45 min. Exposing T4 DNA to 0·20 M-NaOH (in the presence of 1% CH_2O) results in a slightly greater density difference of 0·018 unit. Heating to lower temperatures (83°C and 75°C) results in only a partial displacement in the heavy direction. It seems clear that these are molecules which possess regions of duplex, and regions of single polynucleotide chains. These intermediate bands are broader than those of the native molecules or those of the single polynucleotide chains, because there is some heterogeneity in the fraction of duplex among the various molecules. In each case shown here, a second experiment was performed (not shown) *without* admixing some untreated material. These show no detectable amount of absorbing material in the treated samples remaining in the duplex position. Preparative runs on this material reveal no ^{32}P floating or sinking out of sight as a result of reaction with formaldehyde.

(c) *Sedimentation velocity, and the rate of approach to equilibrium*

The sedimentation coefficients were determined following treatments for various lengths of time at 92°C in the presence of 1% CH_2O. If the treatment did not exceed 4 min, the median S value was about 50 to 55, for material which had about 60 to 62 S in the native condition. When run at pH 12·2 (0·20 M-NaOH) one determination gave 50 S. These values do not conform well with Eigner's calibration (Eigner, 1960) for "denatured" DNA which would extrapolate to an S of about 80 for a molecular weight of about 30×10^6. It would appear that the CH_2O-reacted polynucleotide chains are not equivalent to his untreated preparations, or that these determinations are low for some reason. One should not put too much weight on these values for reasons mentioned previously (Thomas & Berns, 1961).

If one examines the band widths of native and CH_2O-reacted DNA as equilibrium is approached, one notes that the formaldehyde-reacted polynucleotide chains approach equilibrium more rapidly than the native molecules do. This could only happen if the sedimentation coefficient of the singles in CsCl were *greater* than that of the doubles (Meselson, personal communication, 1960). Thus, the S value of the singles appears to be greater than that of the doubles, whatever that value may be.

(d) *Stability of polynucleotide chains under the conditions of treatment*

Samples of DNA at approximately 80 μg/ml. were placed in an Ostwald-Fenske viscometer and held under the following conditions: (1) 92°C with 1% CH_2O; (2) 45°C with 15% CH_2O; and (3) 0·20 M-NaOH with 1% CH_2O at 25°C. In each case, the solvent was 0·10 M-NaCl, 0·032 M-PO_4. In Fig. 4 the results of this experiment are given by plotting the reciprocal of the reduced specific viscosity as a function of time. If we assume that the reduced specific viscosity is roughly proportional to molecular weight (which is not quite true (Eigner, 1960)) and that the fragmentation of the polynucleotide chains takes place randomly, then the results of these experiments may be reported very simply (Charlesby, 1954). At 92°C apparently one molecular break occurs every 10 min.† In 0·20 M-NaOH it requires about 40 hr to do the same, and the 15% CH_2O treatment at 45°C showed no detectable change up to 20 hr.

† Preliminary sedimentation velocity experiments on samples heated for various times show a decrease in S, suggesting that somewhat less than one bond per molecule is broken in 10 min.

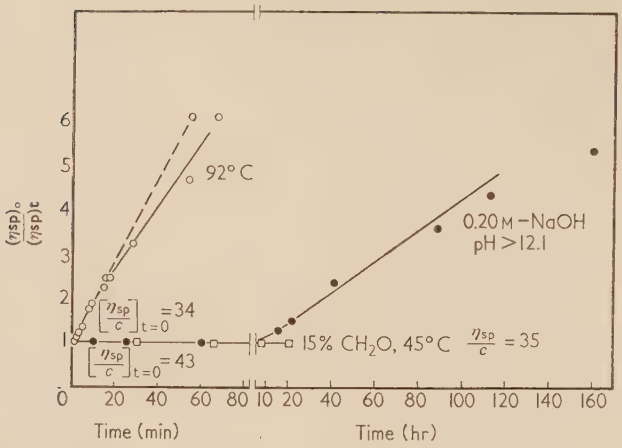


FIG. 4. The relative stability of polynucleotide chains to various treatments. In this figure the ratio of the specific viscosity at the start of the experiment to the same quantity measured at various times is plotted against time. The increase in this ratio should be qualitatively proportional to the number of random breaks in the polynucleotide strands.

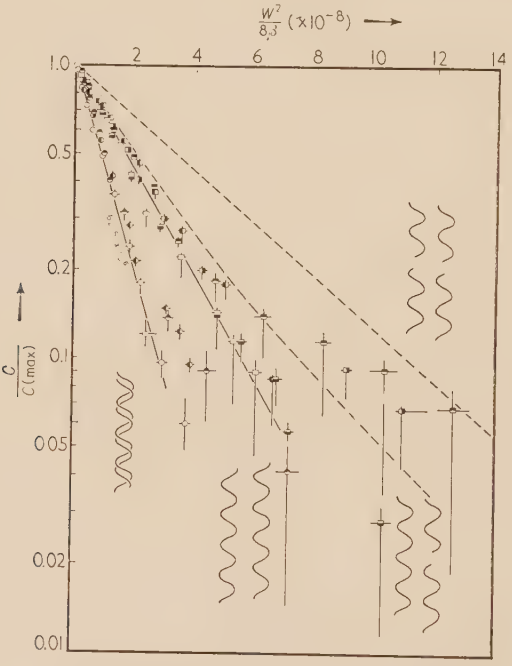


FIG. 5. Comparison of native phenol-extracted DNA with formaldehyde-reacted polynucleotide chains showing band height *vs* (band width)$^2/8\beta$. Density centrifugation of native and denatured T2 and T4 DNA extracted with phenol.

Symbol	Run	Source	C(max)μg/ml.	$b\gamma$	ρ (g/cm³)	Rev/min	$\beta \times 10^4$	
					Native			
○	168	T2H	12·0	6·3	1·693	31,410	0·98	
◓	174	T4D	21·3	21·25	1·698	31,410	1·12	
◐	184	T2H	13·0	23·0	1·691	31,410	1·09	
					Denatured			Method
◧	136	T2H	150·0	3·16	1·715	31,410	1·0	92°C
◨	136	T2H	60·0	6·99	1·714	27,690	1·33	92°C
□	168	T2H	12·2	6·3	1·706	31,410	0·916	92°C
◓	174	T4D	9·8	21·25	1·714	31,410	1·04	(OH^-)
◒	181†	T4D	7·0	19·6	1·712	31,410	1·04	(OH^-)

† DNA produced by osmotic shock.

Thus the most severe treatment (3·0 min at 92°C) might be expected to break a single bond in 30% of the polynucleotide chains.

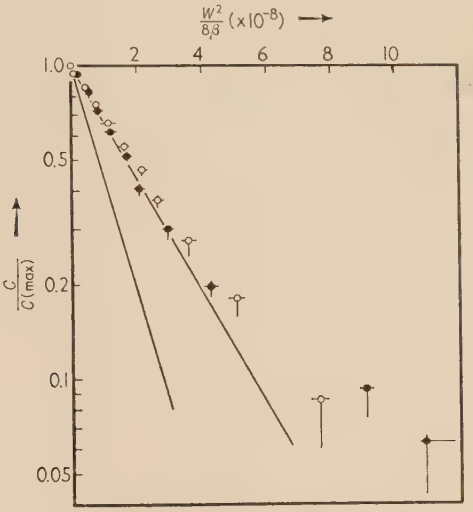


Fig. 6. Plot of band height *vs* (band width)²/8β. Density centrifugation of (OH^-) denatured DNA at two rotor speeds.

○, 31,410 rev/min
●, 27,690 rev/min

(e) *Molecular weight analysis of native DNA and CH_2O-reacted polynucleotide chains*

Experiments such as are shown in Fig. 3 allow one to calculate the molecular weight for both of the bands. All the experiments that have been done are shown in Fig. 5. In this graph the log of the height of the band is plotted against $W^2/8\beta$ where W is the width of the symmetrical band corresponding to the height. The constant β includes all of the constants and variables which do not depend on molecular weight.† By plotting in this way, a number of different experiments may be compared in detail. The points following the steepest curve are for the untreated, undenatured material. As mentioned before, these preparations are compatible with a molecular weight of 62×10^6 (Thomas & Berns, 1961). The formaldehyde-reacted polynucleotide chains display broader bands which correspond to a molecular weight of 31×10^6, exactly half of the undenatured material. On the same graph there are dotted curves corresponding to the expected distributions for molecules containing one or two polynucleotide chain interruptions located near the center. This will be discussed later.

Figure 6 shows the results of banding polynucleotide chains at two different rotor speeds. Changing the rotor speed alters the band width and the concentration of polymer at the band center. The points fall on the same curve, a fact which indicates that there is no observable concentration dependence of the band form (over this concentration range).

† For these calculations we have assumed the simple view of no preferential salt or water binding, electric charge effects, or molecular interactions. The partial specific volume has been equated with the reciprocal of the effective density.

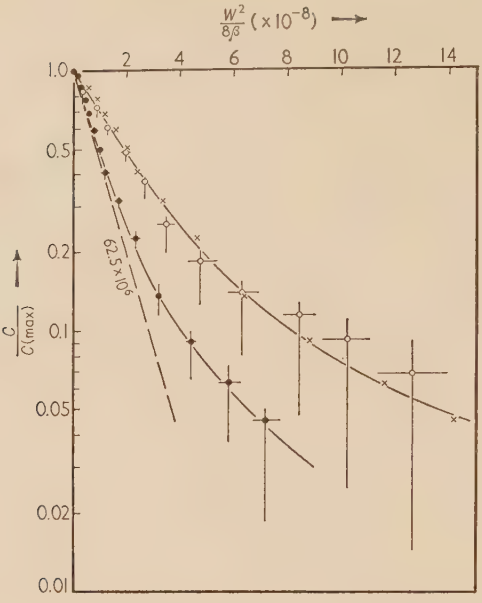


FIG. 7. Plot of band height *vs* (band width)$^2/8\beta$ of CH_2O-reacted polynucleotide chain produced in 0·20 M-NaOH. Density centrifugation of a mixture of native and (OH^-) denatured DNA. The curve through the filled circles (native material) represents the best fit to the data. The curve through the denatured points was derived by doubling the abscissa value for each point in the "native" curve (see Appendix A).

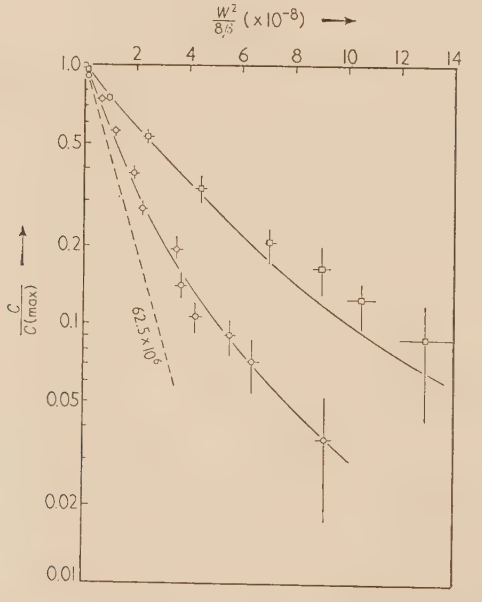


FIG. 8. Plot of band height *vs* (band width)$^2/8\beta$ of shear-degraded DNA and CH_2O-reacted polynucleotide chains. Density centrifugation of native material which has accidentally suffered shear and of denatured DNA (15% CH_2O, 45°C) derived from this material. The native and denatured materials were centrifuged separately.

(f) *Effect of shear*

Figures 7 and 8 show the results of two other experiments on DNA which we presume was accidentally sheared. The molecular weight of the duplex molecules is observably less than 62×10^6 (Fig. 7), or the bands display a pronounced polydispersity effect (Fig. 8). If one assumes that the molecules of DNA making up this collection are double, contain 2 and only 2 polynucleotide chains, then by a theorem outlined in the appendix, the band form of the single polynucleotide chains derived from them should fall on a curve drawn through points which have twice the abscissa of the duplex curve. In other words, without knowing the molecular weight distribution of duplex molecules, it is possible to predict the band form of the singles from the observed band form of the doubles. The curves on the right side of Figs. 7 and 8 are drawn by doubling the abscissa of the curves on the left. The points for the polynucleotide chains fit the curves quite well, a fact which means that the assumption of duplex molecules producing *only* two single polynucleotide chains is a good assumption.

(g) *Effect of dilute DNase*

In order to be sure that no interruptions exist in the native DNA molecule, it is important to demonstrate that one could detect a few interruptions if they did exist.

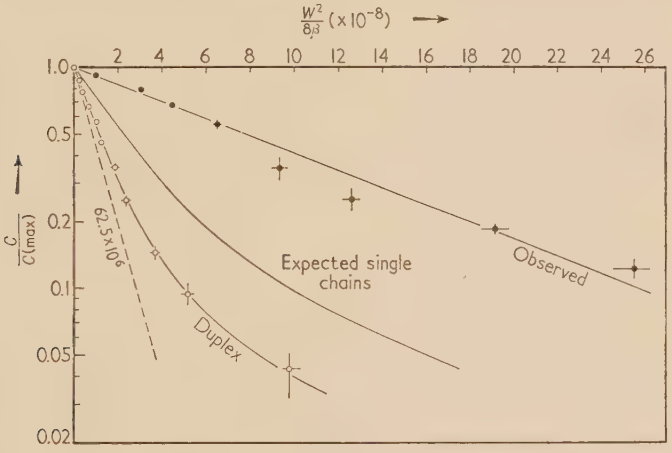


FIG. 9. Effect of DNase treatment on band height *vs* (band width)^2/8β. Density centrifugation of a mixture of native DNA which has suffered damage from dilute DNase and denatured DNA derived from it.

With this in mind, pancreatic DNase (Worthington Co., once crystallized) was added so that there was about one DNase molecule present for every 10 DNA molecules. Under these conditions one can estimate that 5 bonds per polynucleotide chain will be broken at the end of a 6 min period.† The DNase reaction was arrested by the addition of 1% CH_2O and a sample was removed and heated to 92°C for 3 min. Thus in Fig. 9 one sees 2 bands corresponding to duplex and single polynucleotide chains, yet *both* have been reacted with DNase. Now the singles have a much broader band,

† This is based on the assumption that in order to reduce the weight average molecular weight of a DNA molecule (of 62×10^6) by a factor of two, about 2000 ester links must be broken (Thomas, 1956). Under the same experimental conditions it was found that it required 100-fold higher concentration of enzyme to reduce the specific viscosity by a factor of two in 12 min. Since the rate of the reaction appears to be first order with enzyme concentration, one can estimate that approximately 20 bonds will be broken in 12 min (or 10 bonds in 6 min).

and the points fall well outside the expected curve for a duplex producing only 2 polynucleotide chains. The weight average molecular weight is about fourfold lower than expectation. This would be compatible with the existence of an average of 3 random interruptions per polynucleotide chain, or 6 interruptions per molecule of 62×10^6. This is not far from the number (10) of lesions predicted on the basis of enzyme kinetics.

The important thing here is not this agreement but the fact that, when very few breaks are made, the band form of the single is altered as expected.

4. Discussion

The conclusions from this work are fairly simple: that the DNA molecules liberated from phage by phenol extraction are simple duplexes of polynucleotide chains without interruptions. There might be special non-nucleotide linkages, but these are not broken by any of these treatments.

As mentioned in the introduction, in order to establish this fact, it is first necessary to demonstrate that all the hydrogen bonds have been broken, and that none have reformed. On the other hand, one must be sure that no crosslinking is taking place. It seems likely that all the hydrogen bonds have been broken, and none have reformed, because: (1) the optical density has risen to its maximum value and does not decrease on standing; (2) three different methods of denaturation all lead to the same (or nearly the same) increase in density, and give bands of the same molecular weight, while any density heterogeneity (corresponding to heterogeneity in duplex content) would cause the measured molecular weights to be too small or the bands too broad; (3) when breaks are artificially introduced by DNase, the bands do broaden as expected, whereas if they were still associated by hydrogen bonding, they would not.

It appears unlikely that any spurious crosslinking is taking place, and in particular crosslinking mediated by formaldehyde, because: (1) the constancy of band form in spite of the method of preparation—this includes preparing the polynucleotide chains with 15% CH_2O, as well as 1% CH_2O; (2) lower molecular weight duplex material results in lower molecular weight polynucleotide chains, by just a factor of two; (3) DNase results in just the expected broadening of the band; (4) change of rotor speed does not alter the band form (after correction for the new constants); and (5) mixed native and denatured molecules band independently without associating to form species of intermediate density.

If one accepts the technique as being adequate for looking for interruptions, the next question is the resolution of the method. In Fig. 5 the expected distribution for a molecule possessing a single interruption near the center and a similar distribution for a molecule possessing two (non-coincident) interruptions near the center have been plotted. These curves are drawn with the aid of equations (4) and (5) in the previous paper (Thomas & Berns, 1961). The data do not agree with the idea of two breaks, and probably do not agree with the idea of one break. These curves have been drawn assuming *no* polydispersity in molecular weight. Polydispersity causes a "tailing off" of *both* duplex and singles, while a single break case would be expected to cause a tailing off of the singles only. Since none, or very few, of our duplex preparations are totally free of this polydispersity effect, one would expect some tailing in the singles, corresponding to the tailing in the doubles. With this fact in mind, it is clear that the data in Fig. 5 are in better agreement with expectation for two continuous polynucleotide chains, rather than with the model possessing one break near the middle.

Other models with a break near the end of the molecule become increasingly harder to rule out. Thus, in conclusion: the polynucleotide chains are probably continuous, but there is a small chance of there being just one break.

If one accepts this conclusion, there are some implications to phage genetics that should be mentioned. There now seems to be general agreement that the total DNA content of the parental infecting T2 or T4 phage particle is broken down into smaller polynucleotide segments and distributed to the progeny bacteriophage particles (see for example Stent, 1959, p. 268). There is still some disagreement about the maximum (and minimum) sizes of these parental polynucleotide segments. No experiments with T2 or T4 have led to the conclusion that the parental phage DNA is uniquely broken down into 4 (or fewer) segments and distributed to only 4 progeny phage particles; all experiments indicate that the dispersion is greater. Thus, it would appear that links, which we have not broken, are broken during intracellular growth. Likewise, following the dispersal of the parental DNA, the transferred polynucleotide segment is united into the DNA molecules of the offspring phage DNA by links, which we have not been able to break selectively. Thus these conclusions would seem to imply a breaking and subsequent rehealing of the polynucleotide chains during the intracellular growth period. If this is so, it would be of importance to learn whether this breaking takes place at specific sites, or whether it can take place anywhere along the molecule.

The authors wish to thank Drs. Michael Beer and F. D. Carlson for many helpful discussions. Dr. Sigmund Suskind also supplied encouragement and advice as this work developed. Again we would like to thank Dr. A. D. Hershey for supplying us with column-characterized DNA. We are indebted to Dr. Laurence Grossman for pertinent suggestions and the results, prior to publication, of his examination of the addition of CH_2O to nucleotides. This work was supported by the U.S. Atomic Energy Commission Contract AT(30–1)–2119.

REFERENCES

Charlesby, A. (1954). *Proc. Roy. Soc.* A, **224**, 120.
Doty, P., Marmur, J., Eigner, J. & Schildkraut, C. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 461.
Eigner, J. (1960). Thesis, Harvard University, p. 179.
Grossman, L., Levine, S. S. & Allison, W. S. (1961). *J. Mol. Biol.* **3**, 47.
Haselkorn, R. (1959). Thesis, Harvard University.
Levinthal, C. (1959). In *The Viruses*, ed. by F. M. Burnet & W. M. Stanley, Volume II, p. 281. New York: Academic Press.
Marmur, J. & Doty P. (1959). *Nature*, **183**, 1427.
Stent, G. S. (1959). In *The Viruses*, ed. by F. M. Burnet & W. M. Stanley, Volume II. New York: Academic Press.
Thomas, C. A., Jr. (1956). *J. Amer. Chem. Soc.* **78**, 1861.
Thomas, C. A., Jr. (1959). *J. Gen. Physiol.* **42**, 503.
Thomas, C. A., Jr. & Berns, K. I. (1961). *J. Mol. Biol.* **3**, 277.

Appendix A

In this discussion we would like to see if there is any simple relation between the band form of a polydisperse collection of duplex DNA molecules and that of the single polynucleotide strands derived from the duplexes. To do this one may proceed

by writing the general expression for the band form for both single and double forms, and then seeking a simple relation between the independent variables by equating the value of the function.

From equation (4) in the previous paper (Thomas & Berns, 1961) one may write:

$$(C/C_0)_s = \sum_i f_i(s) \exp(-W_s^2 M_i^s/8\beta)$$

$$(C/C_0)_d = \sum_i f_i(d) \exp(-W_d^2 M_i^d/8\beta)$$

to describe the band forms as plotted in Figs. 5 through 9. The quantity $f_i(s)$ is the weighted abundance of the single chains given by the expression:

$$f_i(s) = \frac{w_i^s \sqrt{M_i^s}}{\sum_i w_i^s \sqrt{M_i^s}}$$

where w_i^s and M_i^s are the weight fraction and the molecular weight of the ith molecular weight species. There is an equivalent expression for the double chained form.

$$f_i(d) = \frac{w_i^d \sqrt{M_i^d}}{\sum_i w_i^d \sqrt{M_i^d}}$$

Because each molecule is pictured as separating into *two* polynucleotide chains, $M_i^d = 2M_i^s = 2M_i$. The weight fraction of the ith class is not changed on chain separation, thus $w_i^s = w_i^d$. The result of these assumptions is that

$$f_i(s) = f_i(d) = f_i.$$

Now by equating $(C/C_0)_d$ and $(C/C_0)_s$ one may inquire if there is any simple relation between the band widths for every value of C/C_0.

$$\sum_i f_i \exp(-W_d^2 2M_i/8\beta) = \sum_i f_i \exp(-W_s^2 M_i/8\beta)$$

Since some of the f_i values are greater than zero, and none of them are negative, the only way for this equality to hold is for $2W_d^2 = W_s^2$. This can be easily seen by writing down the terms of these sums and collecting.

Therefore, irrespective of the actual value of the f_i's, the (band width)^2 for the singles is exactly twice that of the doubles measured at any value of C/C_0.

J. Mol. Biol. (1961) **3**, 301–317

Infrared Studies of Molecular Configurations of DNA

E. M. Bradbury,† W. C. Price and G. R. Wilkinson
Wheatstone Physics Laboratory, King's College, Strand, London W.C.2, England

(*Received 31 October 1960*)

The experimental techniques involved in an infrared study of films of biological polymers, whose configurations depend on the water content of the specimen, are described in detail. Polarized infrared spectra of oriented sheets of the sodium and lithium salts of DNA, in both hydrogenated and deuterated states, have been recorded over a range of humidities between 0 and 94% r.h. (relative humidity). For sodium DNA two forms have been shown to exist, one at humidities greater than 90% r.h. where the bases are perpendicular to the helix axis (*B* form) and the other at humidities between 70 and 80% r.h. where the bases have tilted by an angle not less than 13° to the normal to the helix axis (*A* form). For lithium DNA the *B* form is found to exist for humidities greater than 66% r.h. and a second form, different from the *A* form of NaDNA, exists between 44 and 56% r.h. where the bases are tilted by about 4° from the perpendicular (*C* form). The orientation of the phosphate groups, inferred from the dichroic effects of the phosphate bands, is also discussed.

The deuteration rates of both salts of DNA have been measured and found to be extremely fast—less than two minutes for complete deuteration.

1. Introduction

Our knowledge of the major features of the three-dimensional configuration of DNA is based upon information derived from chemical and X-ray diffraction studies which has been correlated by the use of molecular models. Infrared spectroscopy is also capable of providing data of considerable structural value, and measurements of the absorption of polarized infrared radiation give additional and in some cases more accurate information about the orientation of certain groups within the molecule. Whereas X-ray diffraction data refer in the main to the more crystalline parts of a specimen, it is important to note that the infrared spectrum is concerned with the whole of the material.

The molecular configurations of the alkali salts of DNA have been shown to be dependent upon the water content of the film and hence on the relative humidity (r.h.) of the air surrounding the specimen. They depend also upon the nature of the alkali metal. In the case of sodium deoxyribonucleate (NaDNA), Franklin & Gosling (1953) have shown the existence of two configurations, *A* and *B*, at 75% and 92% r.h. respectively. In the *B* configuration the bases lie perpendicular to the helix axis while in the *A* configuration they are tilted by about 20° away from this position (Langridge *et al.*, 1957). Lithium deoxyribonucleate (LiDNA) also exists in two different configurations *viz.* the *B* form, essentially the same as the *B* form of NaDNA at humidities of 66% r.h. and above (Langridge *et al.*, 1957), and the *C* configuration for humidities of 44% r.h., in which the bases are tilted by about 5° (Marvin, Spencer, Wilkins & Hamilton, 1958). X-ray studies of these configurations (Wilkins, Stokes, Wilson & Seeds, 1953; Franklin & Gosling, 1953; Feughelman *et al.*, 1955) have shown

† Now at Courtaulds Research Laboratory, Maidenhead, England.

that the scheme proposed for the structure of DNA by Watson & Crick (1953; Crick & Watson, 1954) is essentially correct. The details of the configuration have, however, been modified considerably to obtain agreement with the X-ray data for the various forms of DNA (Feughelman *et al.*, 1955; Langridge *et al.*, 1957, 1960*a*,*b*). As already mentioned an important distinction between the configurations *A*, *B* and *C* is the difference in the angles which the plane of the base pairs makes with the helix axis. This being so it should be possible to detect these different configurations by polarized infrared studies on oriented films of the salts of DNA. Similar studies would also be expected to help in deciding the orientation of other specific groups in the molecule.

The first infrared studies made on NaDNA were by Blout & Fields (1948, 1949). These, however, were on unoriented films with unpolarized infrared radiation. NaDNA in solution has also been studied by Blout & Lenormant (1953) and by Lenormant & de Lozé (1955). Although this work is important in obtaining information about NaDNA in an unoriented state or in solution it has not been correlated in detail with any of the proposed structures and does not give much information on the configuration of NaDNA. Fraser & Fraser (1951; see also Fraser, 1951) were the first to study oriented films of NaDNA using polarized radiation. Their work, however, preceded the X-ray studies of Franklin & Gosling (1953) and the detailed studies of Wilkins and his group. Thus they did not take into account the importance of the water content of the film in determining the crystallinity and configuration of NaDNA; they discussed only the spectrum recorded at room humidity. The results obtained from their study were correlated with an earlier proposed structure by Furberg (1949) which is now known to be incorrect. The low perpendicular dichroism obtained by Fraser & Fraser (1951) for the in-plane stretching vibrations of the double bonds in the bases supported the earlier proposal by Astbury (1947), based on X-ray data, that the bases were perpendicular to the fibre axis.

Sutherland (1955) and Tsuboi (Sutherland & Tsuboi, 1957) have made detailed infrared studies on oriented films of NaDNA using polarized radiation. Spectra were recorded at twelve different humidities between 0 and 100% r.h. in both hydrated and deuterated states of the specimen. They correlated their results with the modified Watson & Crick structure for DNA (Feughelman *et al.*, 1955) but were unable to demonstrate the existence of the two crystalline forms, *A* and *B*, which occur on variation of the humidity of the air around the specimen. However, as they did not use a reflecting microscope and had to work on relatively large areas of specimen, they could not study the more highly ordered regions which sometimes occur only over small areas of film.

All the above studies have been made on NaDNA, no previous studies having been made on the lithium salt.

Some discussion of the fundamentals of infrared spectroscopy and its application to structure determination is necessary as a preliminary to discussing the details of spectroscopic studies on DNA (see also Bamford, Elliot & Hanby, 1956). In the case of simple molecules in the gas phase the different normal modes in which it is possible for the molecules to vibrate are readily determined by means of group theory and, in the case of molecules with high symmetry, the direction of the transition moments with respect to the molecular axes can also be predicted. It is the direction of the transition moment which determines the direction of maximum absorption for any particular band; this occurs when the electric vector of the radiation is parallel to the transition moment (i.e. to the direction of the change in dipole moment arising from the movement of the atoms). For liquids and solids the theory is much less simple

and the vibrational energy levels are essentially properties of the whole system rather than of individual groups of atoms. Despite this, much progress has been made in the interpretation of the infrared spectra of molecular crystals, including polymers, on account of the fact that the frequencies of the various chemical groups are in many cases independent of their environment. Also it transpires that for some vibrational units the direction of the transition moments associated with the stretching vibrations are not very different from the directions of the chemical bonds, though there are some exceptions, e.g. for C=O groups differences up to 20° between the bond and transition moment direction have been observed. Hence the vibrations of some specific groups may be readily identified from their characteristic frequencies and the direction of a bond may be determined from a study with polarized radiation of the anisotropy of its associated absorption. In the case of many absorption bands, however, especially the low frequency ones, it is not possible to interpret the absorptions simply as being due to the vibration of two atoms as they frequently involve the complex motion of a large number of atoms, and the exact direction of the transition moment associated with the motion is uncertain. Often the directions of transition moments may be altered owing to symmetry of the crystal packing and to interactions between similar neighbouring vibrating groups. Hence comparison with the simple oriented gas model (i.e. with isolated individual units) is not always a valid one, and it is important to stress that caution is necessary in the interpretation of infrared spectra because of the interaction of neighbouring units.

The direction of the transition moments of an absorption band may be obtained from measurements of its dichroic ratio, which may be defined as

$$\frac{\int_{\text{Band}} \alpha_{\perp} d\nu}{\int_{\text{Band}} \alpha_{\parallel} d\nu}$$

where $$\alpha_{\perp} = \log \left(\frac{I_0}{I}\right)_{\perp} \quad \text{and} \quad \alpha_{\parallel} = \log \left(\frac{I_0}{I}\right)_{\parallel}$$

$\frac{I_0}{I}$ is the reciprocal transmission at frequency ν, measured either perpendicular or parallel to the fibre axis, the integration being over the whole absorption band. If the absorption is symmetrical, the maximum absorption $\alpha^{\max}$ is directly proportional to $\int \alpha . d\nu$ and hence the dichroic ratios may be calculated from the simpler expression $\left(\frac{\alpha_{\perp}^{\max}}{\alpha_{\parallel}^{\max}}\right)$. This approximation has been adopted in the present studies.

One of the most valuable uses of infrared spectroscopy in biophysics is for the measurement of the rate of deuterium-hydrogen exchange which occurs when a biological sample is exposed to D_2O vapour. The measurements of the intensities of the absorption bands associated with the NH, OH, ND and OD vibrations permit the relative amounts of any one species exchanged to be determined at any state in the process.

2. Experimental

The following infrared studies on both Na and LiDNA have been made on the specimens prepared by Dr. L. D. Hamilton and used by Dr. M. H. F. Wilkins and his group for their X-ray studies. The spectra were recorded with a double-beam microspectrometer (Ford, 1956; Ford, Price, Seeds & Wilkinson, 1958) using a Grubb-Parsons S_3 monochromator with NaCl and CaF_2 prisms. A 0·72 N.A. reflecting microscope is placed before

the monochromator, and double-beam operation is achieved by moving the sample in and out of the beam 4 times a second. The reference signal (I_0) and sample signal (I) are separated by mechanical rectifiers and fed to a Honeywell Brown recorder which automatically plots the percentage transmission as a function of frequency. The slit width of the monochromator is operated by a Velodyne servomotor so as to obtain a constant I_0 signal. The design of the reflecting microscope is such as to make it particularly suitable for examining fibres and small areas of thin films. The smallest area of film that could be examined in the reflecting microscope was 0·1 mm × 0·01 mm. Although the cone of radiation focused on the specimen was highly convergent, the correction for this has been calculated (Fraser, 1951) and shown to be negligible for films and for fibres where the dichroic ratios are low. The infrared radiation was polarized with a selenium film transmission polarizer (Elliott, Ambrose & Temple, 1948) located behind the exit slit. Since many spectra have to be recorded from the same area of film under varying conditions it was necessary for identifying marks on the film to be aligned with the slit. For this purpose a small removable viewing prism was fitted behind the entrance slit so that the image of the specimen on the slit could be seen through the prism with a low-power microscope.

(a) *Specimen preparation*

Specimens were prepared as thin films on plates of silver chloride or barium fluoride, the latter on account of its greater hardness being superior for frequencies down to 800 cm^{-1}. Orientation of the films was obtained by unidirectional shearing of very viscous aqueous solutions with a carefully ground horn spatula until the films were dry. These films showed high negative birefringence when viewed between crossed nicols in a polarizing microscope. Small areas of film showed uniform interference colour of high extinction and, although this is not necessarily a criterion for high orientation, it was found that these areas gave the highest dichroic ratios. Areas of poor extinction always gave low dichroic ratios and hence were avoided.

Table 1

Relative humidity	Salt	
94%	Sodium sulphate	Na_2SO_4
86%	Potassium chloride	KCl
76%	Sodium chloride	NaCl
66%	Sodium nitrate	$NaNO_2$
56%	Sodium bromide	NaBr
44%	Potassium carbonate	K_2CO_3
32%	Calcium chloride	$CaCl_2$
20%	Potassium acetate	$CH_3.COOK$
15%	Lithium chloride	LiCl
5%	Sodium hydroxide	NaOH
0%	Phosphorous pentoxide	P_2O_5

The data on the above salts are reproduced from "Laboratory and Workshop Notes" edited by R. Lang.

(b) *Humidity control*

As the configuration of Na and LiDNA molecules was known to be dependent on the amount of water present in the film, a small constant humidity cell was constructed for use with the reflecting microscope. The cell consisted of two halves; the upper half was a metal annulus which carried the specimen, and the lower half carried a window of barium fluoride or silver chloride sealed in with stopcock grease. The upper portion was screwed on to an O-ring held in a groove in the lower part of the cell, sealing the specimen film on the inside. Humidity control was effected by placing saturated solutions of various salts, with excess solid, in a well in the lower half of the cell. Humidities were chosen (Table 1) such that the controlling salts had a constant temperature coefficient between

20° and 25°C for their relative humidity (O'Brien, 1948). The spectra were recorded in this temperature range. To prevent wastage of deuterium oxide in deuteration experiments, anhydrous salts were used wherever hydrates might be formed. The water content was also measured experimentally by observing the integrated band area of the absorptions due to the OH and OD stretching vibrations.

(c) *Deuteration*

As the salts of DNA have a high water content the important 3300 cm^{-1} and 1650 cm^{-1} regions are normally masked by strong water bands. Replacement of this water by deuterium oxide leaves these regions clear since the D_2O bands occur at lower frequencies. Absorption bands which arise from vibrations of groups containing replaceable hydrogens are also shifted to lower frequencies and, by observing the rate of change in intensity of these bands with time after the deuterating agent has been added, the rate of deuteration of the specimen can be measured.

Deuteration was effected by exchange with solid films from atmospheres containing constant humidities of D_2O. The humidity was controlled by making solutions of the salts mentioned in the previous section with 99·78% deuterium oxide. The small cell used for humidity control was ideal for deuteration purposes. All operations were carried out in a dry box flushed with air dried by passing over activated alumina. The cell was dried and a saturated deuterium oxide solution was placed in the well in the lower half. It was then sealed with the specimen in the upper half and placed in the microscope. By observing the rate at which the intensity of absorption of the 3400 cm^{-1} region decreased with time the deuteration of the specimen could be followed. When the specimen was completely deuterated—which in the case of salts of DNA took less than 2 min (i.e. the time it took to seal the cell and mount it in the microscope)—the relative humidity (D_2O) could then be varied by changing the saturated salt solutions of deuterium oxide in the well in the bottom of the cell. The constant-humidity cell was tested for effectiveness by sealing a deuterated film inside and leaving it in open atmosphere for 7 days; at the end of this time the specimen was still completely deuterated and no back exchange had taken place.

3. Results and Interpretations

Infrared spectra of Li and NaDNA were recorded at different humidities between 5 and 100% r.h. Two of these spectra for LiDNA are shown in Fig. 1 where it can be seen that the changes on varying the water content of the film are quite remarkable —similar changes have also been pointed out by Sutherland & Tsuboi (1957) for NaDNA. In changing from high to low humidity the dichroism of many bands decreases and frequency shifts occur, which for some bands are quite large. At high humidities too the bands in the region 750 to 1200 cm^{-1} are sharper and better defined. This indicates that the crystallinity of the specimen improves with water uptake. In following the changes that take place in the orientation of the bases with change of humidity the main region of interest is between 1750 and 1550 cm^{-1} where the in-plane vibrations of the bases are found. Deuteration shifts the water band which normally masks this region and also the band which arises from the deformation vibration of the NH_2 groups in the bases. The remaining double-bond stretching bands now show high dichroism, as shown in Fig. 2, and any changes in their orientation can be followed by measuring the dichroic ratio as a function of water content of the film. On lowering the D_2O content of the film the general changes in the spectrum are similar to those observed for the hydrated films.

(a) *The spectrum of LiDNA*

The spectrum of LiDNA is similar to that of NaDNA and shows three regions of strong absorption (1) between 3500 and 2500 cm^{-1}, (2) between 1750 and 1550 cm^{-1}

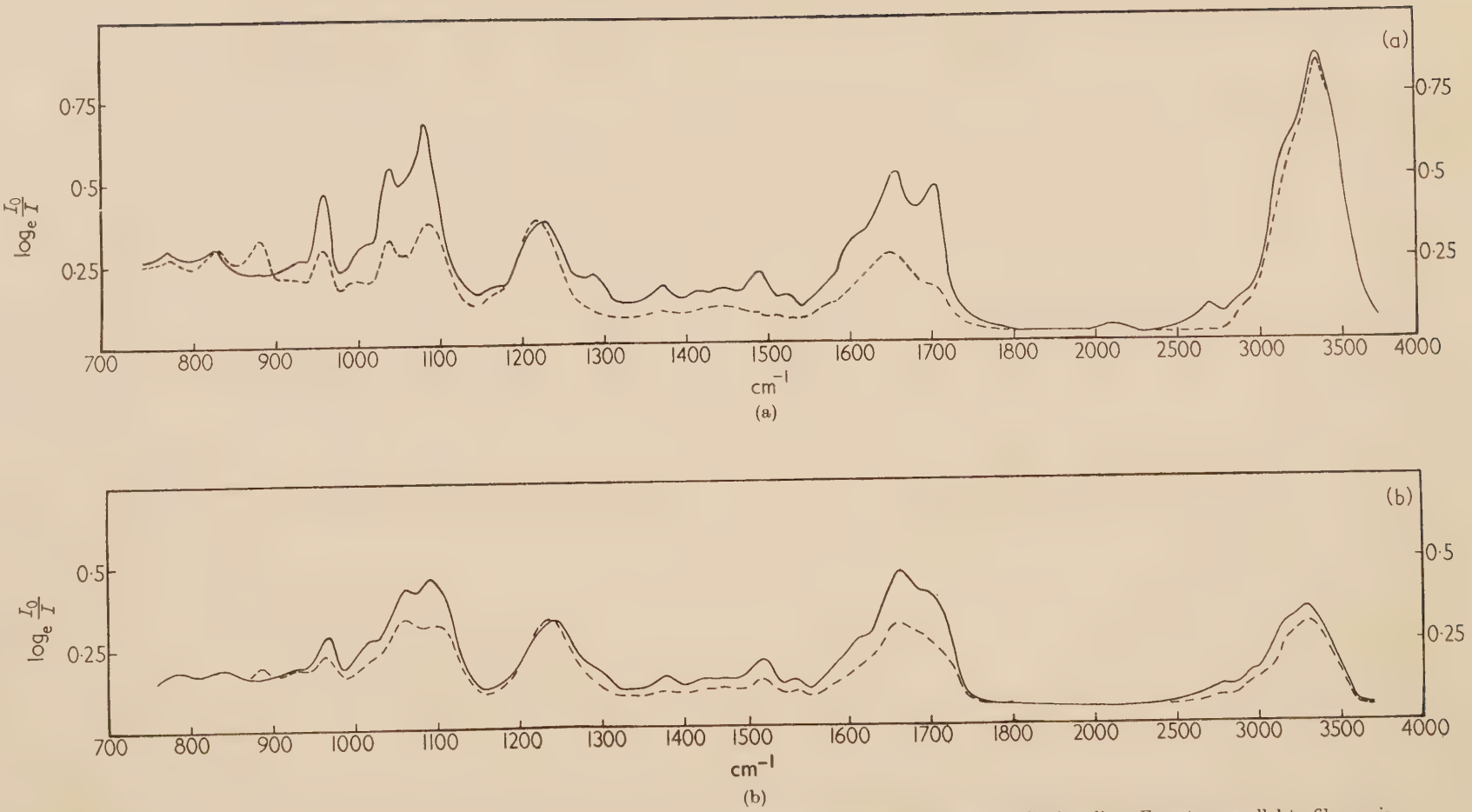


FIG. 1. Infrared spectra of LiDNA (a) at 86% r.h.; (b) at 15% r.h. Full line, E vector perpendicular to fibre axis. Broken line, E vector parallel to fibre axis.

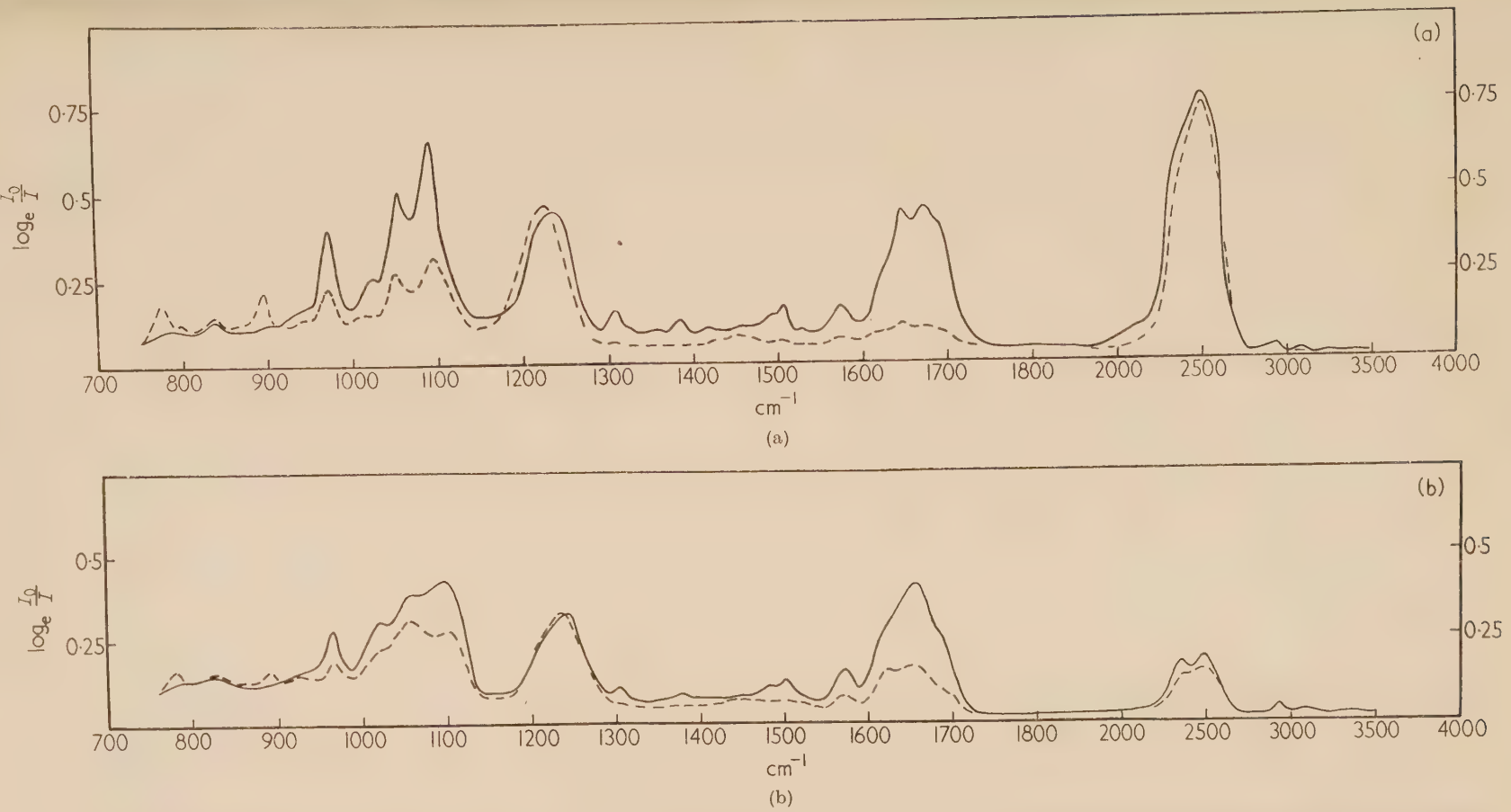


FIG. 2. Infrared spectra of deuterated LiDNA (a) at 86% r.h. D_2O; (b) at 15% r.h. D_2O. Full line, *E* vector perpendicular to fibre axis. Broken line, *E* vector parallel to fibre axis.

and (3) between 1300 and 900 cm^{-1}. These absorptions have been shown to arise from vibrations of specific groups in the DNA molecule; the bases, together with absorbed water, give strong bands in the first two of the above three spectral regions (Sutherland & Tsuboi, 1957; Bradbury, 1959); the sugar units contribute bands to the first and third of the regions; and the phosphate group has very strong bands in the 1300 to 900 cm^{-1} region only (Sutherland & Tsuboi, 1957; Tsuboi, 1957; Bradbury, 1959). The details of many assignments have already been given by Sutherland & Tsuboi and will only be discussed briefly.

(i) *3500 to 2500 cm^{-1} region*

The spectra of Li and NaDNA at high humidities both show a strong band at about 3395 cm^{-1}, with several shoulders on the low frequency side at 3200, 2950 and 2760 cm^{-1}. On drying the film the greater part of the absorption at 3395 cm^{-1} disappears, showing that it is a water band. Deuteration removes the four strong bands leaving weak vibrations at 3095 and 2850 cm^{-1}. The former is probably the ring C—H stretching vibration of the bases and the latter the antisymmetric vibration of the CH_2 groups in the sugar part. The bands that shift can be assigned to a water band and to stretching vibrations of the NH and NH_2 groups in the bases. Unfortunately the deuteration rate is so high that the hydrogen atoms in the NH and NH_2 groups, which take part in the hydrogen-bonding mechanism, exchange before their bands can be recorded unobscured by water absorption. The very small dichroism seen in this region is due to the presence of a large amount of amorphous water absorbed by the film.

(ii) *1750 to 1550 cm^{-1} region*

This region of the spectrum is the most informative about changes of configuration of the bases in both Na and LiDNA during water uptake of the film. At room humidity a broad band is found with its maximum at about 1600 cm^{-1}, and there are several shoulders on the high and low frequency sides. Frequencies and intensities change on varying the relative humidity of the film in both the hydrated and deuterated states. For the hydrated film at high humidity two strong bands are found at 1711 cm^{-1} and 1665 cm^{-1} with shoulders at 1620 and 1580 cm^{-1}. On drying the film the 1711 cm^{-1} band moves progressively to lower frequencies until in the dry state it is found at 1689 cm^{-1}. The bands at 1665, 1620 and 1580 cm^{-1} move by only a few wave numbers to 1666, 1618 and 1582 cm^{-1}. In the dry state the shoulders at 1618 and 1582 cm^{-1} have become more pronounced owing to the removal of water. The whole region shows perpendicular dichroism. At high humidities the band at 1711 cm^{-1} shows a much higher dichroism than that at 1665 cm^{-1} owing to the superposition of bands. As the absorption in the 1660 cm^{-1} region increases with water content it is clear that this added absorption is due to the ν_2 bending vibration of water. The remaining bands can be assigned to double bond (C=O, C=N, C=C) stretching vibrations in the bases together with the deformation vibration of the NH_2 groups.

The effect of deuteration can be seen in Fig. 2, and the removal of the water band and the NH_2 deformation band reveals a much increased perpendicular dichroism. Frequency shifts are also observed on comparison with the hydrated film. Bands are found at 1689, 1672, 1650, 1629 and 1575 cm^{-1} at 86% r.h. D_2O. The band at 1711 cm^{-1} in the hydrated film has moved to a lower frequency on deuteration and is now probably the band at 1689 cm^{-1}. In the bases which have a carbonyl group the band

associated with the stretching vibration of this group is always found at frequencies higher than 1680 cm^{-1}. For guanine it is at 1691 cm^{-1}; in thymine, with two carbonyl groups, the stretching frequencies are at 1748 cm^{-1} and 1685 cm^{-1} (Angell, 1955; Bradbury, 1959). For cytosine the absence of this band, which might be expected at frequencies greater than 1680 cm^{-1}, was taken to indicate that this base was in the enolic form. In cytosine deoxyriboside, where the sugar unit is attached to the N_3 atom, the question of tautomerism does not arise and cytosine in the nucleoside is in the ketonic form; in this case the carbonyl stretching vibration results in a band at 1712 cm^{-1} (Bradbury, 1959). The band at 1711 cm^{-1} in DNA is also probably a carbonyl stretching frequency of one of the bases. The shift from 1711 cm^{-1} to 1689 cm^{-1} on deuteration probably results from exchange of the hydrogens in the NH and NH_2 groups; the vibration of the carbonyl groups is probably not independent of the vibrations of the NH and NH_2 groups. Similar shifts of carbonyl frequencies on deuteration are also noticed in proteins, in collagen (Bradbury, Burge, Randall & Wilkinson, 1958), protamine and histone (Bradbury, 1959) on complete deuteration. Here the carbonyl stretching frequency shifts to a lower value by about 12 cm^{-1}. In acetanilide (Abbott & Elliott, 1956) a reduction of 21 cm^{-1} in the frequency of the carbonyl stretching vibration is observed on deuteration.

If the dichroic ratio $\left(R = \frac{\alpha_{\perp}^{max}}{\alpha_{\parallel}^{max}}\right)$ of the 1672 cm^{-1} band for the deuterated film of NaDNA is plotted against the water content of the film the curve shown in Fig. 3

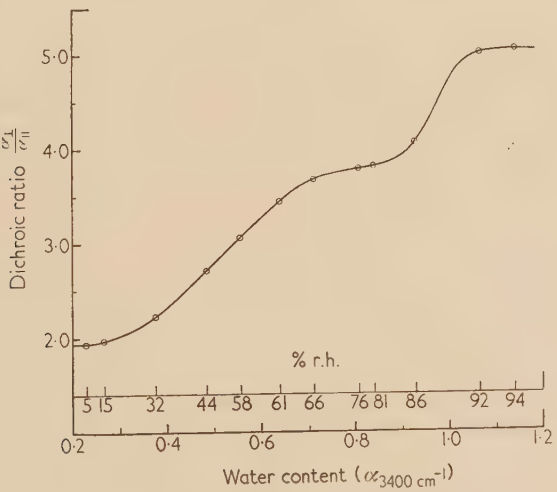


Fig. 3. Plot of the dichroic ratio of in-plane base vibrations for NaDNA against water content (D_2O) of the specimen.

is obtained. 1672 cm^{-1} was selected for the calculation of the dichroic ratio of the in-plane vibrations of the bases because it gave the maximum polarization ratio and because it lies close to the centroid of the base absorptions. Absorption due to other parts of the polymer is negligible in this region. It can be seen that two plateaux exist, one for humidities greater than 90% r.h. and the other for humidities between 70% and 80% r.h. These plateaux must correspond to the two crystalline forms of Franklin & Gosling (1953), the *A* form at 75% r.h. and the *B* form at 92% r.h. In

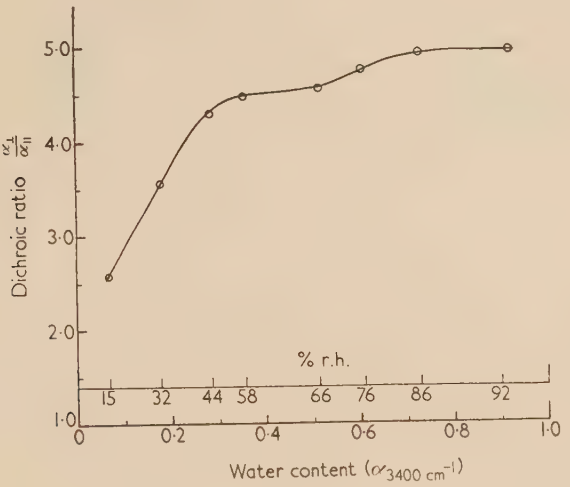
the *B* form the planes of the base pairs are perpendicular to the helix axis while in the *A* form they are tilted through an angle of 20° (Langridge *et al.*, 1957). The dichroic ratio for the in-plane vibrations of the bases of the *B* form is 5·1 : 1 while in the *A* form it is 3·7 to 3·8 : 1. Since X-ray evidence shows that it is impossible by any known method of specimen preparation to produce perfect orientation in a film of NaDNA, the dichroic ratio of 5·1 : 1 probably indicates that the bases are very nearly perpendicular to the helix axis. However, if it is assumed that the bases are perpendicular to the helix axis of each molecule, then using the expression (Fraser, 1951)

$$R = \frac{\frac{1}{2}F \sin^2 \theta + \frac{1}{3}(1 - F)}{F \cos^2 \theta + \frac{1}{3}(1 - F)}$$

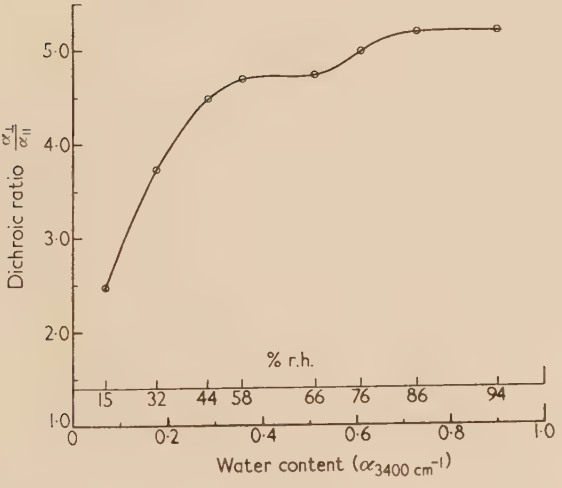
which refers to a model such that F is the fraction of completely ordered molecules, $(1 - F)$ having random orientation, and θ is the angle between the transition moment and the helix axis, then substituting $\theta = 90°$ and $R = 5{\cdot}1$ the value of F is found to be $\frac{8{\cdot}2}{11{\cdot}2}$, i.e. 73% of the molecules are completely ordered. Using this value for F and the dichroic ratio for the plateau corresponding to the *A* form, i.e. 3·8 : 1, θ is found to be 77°; that is to say, the tilt is 13°. This figure must be regarded as a minimum value since the direction of transition moments in the plane of the bases are unknown and it is unlikely that they will be along the direction of maximum tilt. If they were to lie perpendicular to this direction of maximum tilt then they would remain perpendicular to the fibre axis irrespective of the angle through which the bases tilt and the dichroism would remain unaltered. The actual tilt of the bases is most likely to be higher than 13° and the value of 20° determined by X-ray studies is in no way disputed. Although Sutherland & Tsuboi (1957) did not find the definite indication of the two forms of NaDNA shown in Fig. 3, they did find a change in dichroic ratio between 92% and 75% r.h. from 2·7 : 1 to 2·1 : 1. They repeated the calculation described above and the minimum angle the transition moments make with the helix axis for the *A* form turns out to be 75° (i.e. a tilt of 15°). The two values are in close agreement.

For LiDNA similar plots of dichroic ratio against water uptake of the film are shown in Fig. 4(a) and (b). Here again two plateaux are found which probably indicate that two forms exist for LiDNA as well as for NaDNA. However, the difference between the two plateaux is much less marked than for NaDNA and the humidity ranges over which they exist are also different. These two plateaux can be correlated with the two forms, *B* and *C*, of LiDNA. The *B* form (Langridge *et al.*, 1957), found at humidities of 66% r.h. and greater, is the same as the NaDNA *B* form with the bases perpendicular to the helix axis. In the *C* form (Marvin *et al.*, 1958), found at 44% r.h., the planes of the bases are tilted by about 5°. Figure 4(a) shows that for relative humidities of 66% and above the LiDNA is in the *B* form and gives a dichroic ratio of 5·2 : 1. At humidities between 44% and 56% r.h. the dichroic ratio falls to 4·8 : 1, which is indicative of the *C* configuration. Making the same calculation as for NaDNA then $F = 73\%$, and using this value of F together with $R = 4{\cdot}8$ for the *C* form we find that $\theta = 85\frac{1}{2}°$, indicating a minimum angle of tilt of the bases of about $4\frac{1}{2}°$. An error of $\pm$ 0·1 in the measured dichroic ratio leads to an error of 1° in the angle of tilt. From the scatter of measurements of dichroic ratio the probable error was found to be about $\pm$ 0·1 which indicates that the minimum angle of tilt is $4\frac{1}{2} \pm 1°$. Figure 4(b) shows a similar curve with slightly lower dichroic ratios for the two

plateaux (5·0 : 1 and 4·6 : 1). These figures give a minimum angle of tilt which is also about $4\frac{1}{2} \pm 1°$.



(a)



(b)

FIG. 4(a), (b). Plot of dichroic ratio of in-plane base vibrations for LiDNA against the water content (D_2O) of the specimen. The differences between these spectra of different preparations of LiDNA probably arise from variations in chloride content of the specimens (Marvin *et al.*, 1961).

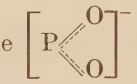
It may be noted that the dichroic ratio change in Figs. 4(a) and 4(b) takes place at about 75% r.h. and this indicates that the transition from *B* to *C* configuration in LiDNA also takes place at about 75% r.h. X-ray diffraction studies (Marvin *et al.*, 1961) show that if the chloride content of LiDNA is such that the crystalline form of the *B* configuration occurs, the transition from *B* to *C* takes place at 50% r.h. or below ; however, when the chloride content is lower, the *B* to *C* transition takes place at 75% r.h., as shown in Fig. 4(a) and 4(b). Evidently the specimens used in these infrared studies had a low chloride content.

(iii) *1550 to 1325 cm^{-1} region*

Seven weak bands, all showing high perpendicular dichroism, are found in this region in the hydrated state. Assignments of these bands are very difficult as it is doubtful whether any one of them results from a simple mode of vibration. The bases and sugar units give rise to many bands here, and as the general level of absorption of DNA in this region is high this would indicate that the peaks arise from the addition of many bands. Deuteration produces little change, a few bands shifting by a very small amount. The band at 1495 cm^{-1} shifts, leaving bands at 1501 and 1484 cm^{-1}. That this region shows high perpendicular dichroism probably indicates that the majority of bands are from in-plane base vibrations.

(iv) *1325 to 1200 cm^{-1} region*

The prominent feature of this region is the strong non-dichroic band found at 1228 cm^{-1} at high humidities, which shifts on drying the film to 1245 cm^{-1}. There are also two weak bands at 1301 and 1284 cm^{-1} which show high perpendicular dichroism. Deuteration does not affect the very strong band but shifts the weak band at 1284 cm^{-1}—the disappearance of this weak band has important applications in the studies of nucleoproteins, an account of which will follow in subsequent papers. A new band also is seen at 1219 cm^{-1}; this is a D_2O band.

It is now generally agreed that the strong band at 1228 cm^{-1} arises from the antisymmetrical stretching vibration of the  group. This assignment is supported by the detailed study made by Tsuboi (1957) and others of the vibrations of the PO_2^- group. In both potassium hypophosphite [KH_2PO_2] and barium dimethylphosphate $\{(CH_3O)_2PO_2^-\}_2Ba^{2+}$ Tsuboi has found a strong band in the 1230 cm^{-1} region which shifts to lower frequencies in solution. This band he has assigned to the antisymmetrical stretching vibration of the PO_2^- group. The spectrum of sodium diethylphosphate $\{(C_2H_5O)_2PO_2\}^-Na^+$ also supports this assignment (Bradbury, 1959). Here a strong band is found at 1242 cm^{-1} for the dry salt deposited as a microcrystalline film on barium fluoride—on raising the humidity of the film the band moves to 1211 cm^{-1} at 75% r.h. This is clearly the same band as in Na and LiDNA.

The behaviour of the 1228 cm^{-1} band on varying the water content of the film is very interesting. In the dry state the perpendicular and parallel components of this band are coincident and show no dichroism. On raising the humidity both components move to lower frequencies and separate until at 94% r.h. the separation is 6 cm^{-1} with the perpendicular component at higher frequencies. For NaDNA the frequencies of the components of the strong band are plotted against the water content of the film in Fig. 5. A double plateau is found which shows that the frequency is constant over the ranges of humidity where the *A* and *B* configurations are observed. It appears that each configuration has a characteristic frequency. The same plot for LiDNA does not show a double plateau; if there is a frequency difference for the two forms *B* and *C*, it is not large enough to be detected. This shift of frequency to lower values as the water content of the film increases is probably due to the binding of water molecules onto the PO_2^- group.

The transition moment of the antisymmetric stretching vibration of the PO_2^- group lies along the O . . . . O line. In the dry film, where there is a dichroic ratio of 2·1 for the in-plane double band vibrations of the bases, the 1246 cm^{-1} band shows

no dichroism and, since the perpendicular and parallel components are coincident, this lack of dichroism indicates that the O . . . . O line lies very close to 55° to the helix axis. On raising the humidity the specimen becomes much more crystalline and the perpendicular and parallel components separate by about 6 cm^{-1} at 94% r.h. This splitting of the band must show the presence of an interaction between vibrations of adjacent phosphate groups which is dependent on the crystallinity of the specimen (Sutherland & Tsuboi, 1957). The fact that both components have the same intensity suggests, as in the dry film, that the O . . . . O line lies very close to 55° to the helix axis.

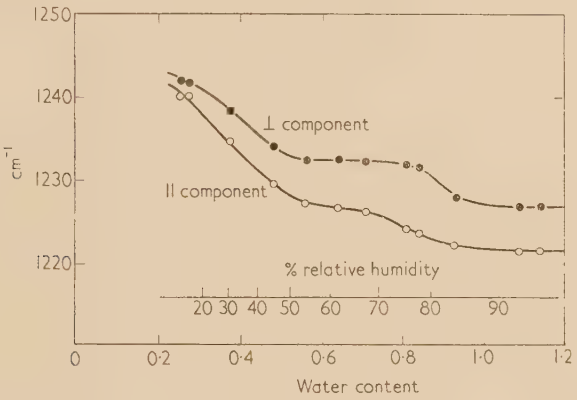


FIG. 5. Plot of the frequencies of the components of the band due to the antisymmetrical stretching vibration of the PO_2^- group, against the water content of the specimen.

(v) *1200 to 900 cm^{-1} region*

The strong bands found in this region can be associated with vibrations of groups in the phosphate and sugar parts of DNA. Bands are found at 1089, 1070, 1053, 1019 and 971 cm^{-1} at high humidities. The 1089 and 1019 cm ^{1} frequencies refer to the perpendicular components of the bands; the parallel components are slightly separated and are found at 1092 and 1014 cm ^{1}. On drying the film, frequency shifts occur in three of the bands and the whole region becomes ill-defined. The bands at 1089 and 1053 cm^{-1} move to higher frequencies, 1096 and 1065 cm^{-1} respectively, which is similar to the shift of the antisymmetric stretching vibration of the PO_2^- group. The band at 971 cm^{-1} moves to a lower frequency, 964 cm^{-1}. Deuteration produces no change in this region.

The strong band at around 1230 cm^{-1} has been assigned to the antisymmetric stretching vibration of the PO_2^- group and, because the symmetric stretching vibration of this group is expected at lower frequencies, it would probably give rise to a band in the 900 to 1250 cm^{-1} region. Tsuboi (1957) has assigned a band at 1048 cm ^{1} in the hypophosphite anion $[H_2PO_2]^-$ to the symmetric stretching vibration of the PO_2^- group and a band in barium dimethylphosphate at 1054 cm ^{1} has been similarly assigned. In sodium diethylphosphate (Bradbury, 1959) a band is found at 1051 cm^{-1} at 75% r.h. which can also be assigned to the symmetric stretching vibration of the PO_2^- group. This band behaves in a similar manner to the band at 1053 cm ^{1} in DNA in that it moves to higher frequencies from 1051 to 1060 cm ^{1} on drying the film. The band in DNA moves from 1053 to 1061 cm ^{1}. From the work by Tsuboi (1957).

Sutherland & Tsuboi have assigned the band at 1053 cm^{-1} in DNA to the symmetric stretching vibration of the PO_2^- group. Absorption bands due to C—O—C vibrations also arise in this region and may contribute to the band at 1054 cm ^{1}. However, if this is so the contribution must be small since the band behaves in a manner identical to the similarly assigned band in the compounds mentioned above where there are no C—O—C groups. This band shows perpendicular dichroism like all the bands in this region. Thus the transition moment of the symmetric stretching vibration of the PO_2^- group, which is along the OPO bisector, must by closely perpendicular to the helix axis.

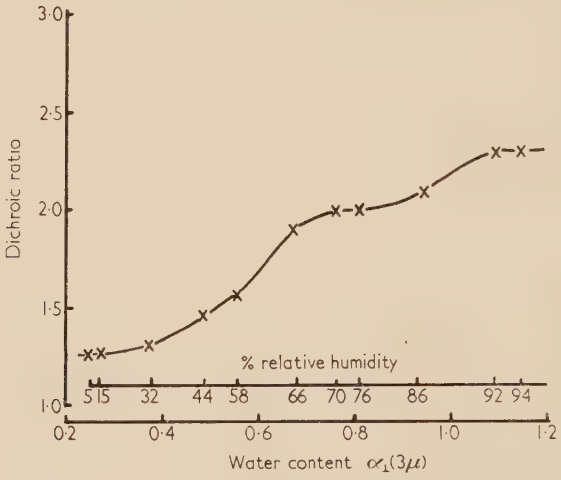


Fig. 6. Plot of the dichroic ratio of the 1053 cm^{-1} ($\perp$) band with intensity due to the symmetrical stretching vibration of the PO_2^- group, against water content of the specimen.

If the dichroic ratio of the 1053 cm ^{1} band in NaDNA is plotted against the water uptake of the film, the curve shown in Fig. 6 is obtained. Here again we observe a double plateau, with dichroic ratios of 2·3 and 2·0 : 1, indicating the presence of the two forms, *B* and *A*, of NaDNA. For LiDNA the dichroic ratio of this band at 86% r.h. is 2·3 : 1. The curve obtained by plotting this against water uptake shows one plateau extending down to 58% r.h. after which the dichroic ratio falls. Two plateaux are not observed, probably because there is very little difference in the orientation of the phosphate group between the *B* and *C* forms. The value of this dichroic ratio for the *B* configuration of both LiDNA and NaDNA, 2·2 to 2·3 : 1, is important since the phosphate region is unobscured by protein bands in the nucleoproteins. It gives an indication of the configuration of DNA in protein—DNA complexes.

The remaining bands in this region arise from other vibrations of the phosphate and sugar parts of the DNA molecule. Bands arising from C—O—P and C—O—C vibrations would also be expected in this region. The band at 1089 cm^{-1} which shifts to 1096 cm^{-1} on drying the specimen is possibly in part connected with a phosphate vibration and may well be attributed to the C—O—P vibration. There is a strong band in sodium diethylphosphate at 1084 cm^{-1} at 75% r.h. which moves to 1095 cm^{-1} on drying, and this could have the same assignment.

(vi) *Below 900 cm^{-1}*

There are several weak bands below 900 cm^{-1}. One at 894 cm $^{-1}$ shows high parallel dichroism and is unaffected by deuteration—this has been assigned to the out-of-plane C—H bending vibration of the C—H groups in the bases (Wilkinson, Price & Bradbury, 1959). Deuteration also results in the appearance of a new weak band of high parallel dichroism at 777 cm^{-1}, the origin of which is obscure.

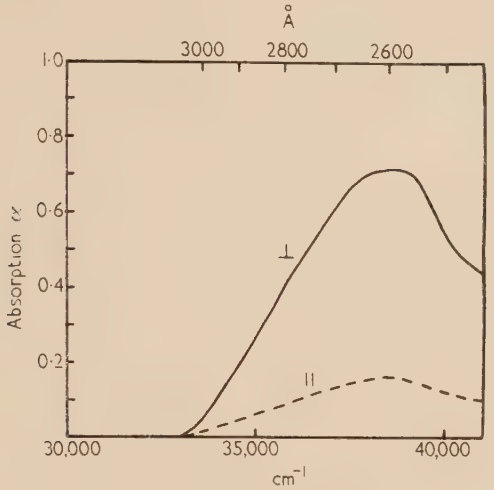


Fig. 7. The ultraviolet absorption spectrum of NaDNA recorded with the electric vector perpendicular and parallel to the polymer axis (Seeds, 1951). Full line, *E* vector perpendicular to fibre axis. Broken line, *E* vector parallel to fibre axis.

(b) *The ultraviolet spectrum of DNA*

The ultraviolet absorption spectrum of DNA has been recorded by Seeds (1951) using polarized radiation. Figure 7 shows the spectrum of an oriented sample at a relative humidity of 90%, recorded with the electric vector of the radiation first parallel and then perpendicular to the polymer axis. The specimen was prepared in a manner similar to that used for the infrared studies; however, the film thickness was much less, a thickness of only about 0·25 micron being used for the determination of the ultraviolet dichroism. The u.v. absorption in the 30,000 to 40,000 cm^{-1} region is due to electronic transitions in the bases of both $n - \pi$ (weak) types and $\pi - \pi$ (strong) types. The latter are known to have transition moments in the planes of the bases, and hence the maximum absorption occurs when the electric vector of the radiation is perpendicular to the polymer axis. The dichroic ratio of 4·7 to 1 obtained from these ultraviolet absorption bands agrees well with the values observed for the dichroic ratio of the infrared absorption bands due to the in-plane vibrations of the bases.

The variation in the intensity of this ultraviolet absorption band is of great value in studying the configuration of DNA in solution. On breaking the twin helical structure the absorption at 37,700 cm^{-1} (2650 Å) increases by about 45%.

4. Discussion

Franklin & Gosling (1953) have shown that NaDNA exists in two crystalline forms, *A* and *B*, according to the humidity of the air around the specimen. The *B* form has

the planes of the base-pairs perpendicular to the fibre axis and is found at humidities greater than 90% r.h., while the *A* form has the base-pairs tilted by 20° and is found at humidities in the region of 75% r.h. The present infrared studies, which have been done on more highly ordered specimens than those used by Tsuboi & Sutherland, confirm the existence of the two crystalline forms over the same ranges of humidity. The dichroic ratio of 5·1 : 1 found for the *B* form shows that the bases are probably closely perpendicular to the helix axis. If it is assumed that the bases are perpendicular to this axis, then the fraction of molecules which are completely oriented, the remainder having random orientation, can be calculated to be 0·73. Since the direction of the transition associated with the in-plane double-bond stretching vibrations of the bases is not known within the plane of the bases, it is only possible to give a lower limit for the tilt of the bases. This is 13° for the *A* form. For LiDNA X-ray studies have shown that two forms, *B* and *C*, again exist. The *B* form is the same as the *B* form in NaDNA with base-pairs perpendicular to the helix axis for humidities of 66% r.h. and above (Langridge *et al.*, 1957). The *C* configuration exists at a humidity of 44% r.h., the base pairs being tilted by 5° (Marvin *et al.*, 1958). Both forms have also been observed in these infrared studies, which are the first to be made on LiDNA. The minimum tilt of the bases has been shown to be $4\frac{1}{2} \pm 1°$ and the fraction of totally oriented molecules in the specimen is evaluated as 0·72. These results agree closely with those from X-ray studies.

The orientation of the PO_2^- group can be determined with reasonable certainty from the infrared studies. These have shown that for the *B* form the O . . . . O line in the PO_2^- group lies at 55° from the helix axis and the OPO bisector is not less than 70° from the fibre axis (calculated from $R = 2{\cdot}3$ and $F = 0{\cdot}73$); as pointed out by Sutherland & Tsuboi (1957), this is different from the orientation of the PO_2^- group in the first structure proposed by Crick & Watson (1954). For the *A* form of NaDNA the OPO bisector is probably rotated by a few degrees (approximately 3°) away from the perpendicular without changing the orientation of the O . . . . O line.

The fast deuteration rate of salts of DNA measured by vapour phase exchange shows that the hydrogen atoms which take part in the hydrogen-bonding mechanism between the bases are completely accessible to protons and possibly to water molecules.

We wish to thank Dr. M. H. F. Wilkins and co-workers for discussion and Dr. L. D. Hamilton for supplying specimens of DNA. One of us (E.M.B.) wishes to thank D.S.I.R. for a maintenance grant.

REFERENCES

Abbott, N. B. & Elliott, A. (1956). *Proc. Roy. Soc.* A. **284**, 247.
Angell, C. (1955). Ph.D. Thesis, Cambridge.
Astbury, W. T. (1947). *Symp. Soc. Exp. Biol.* I. p. 66. Cambridge University Press.
Bamford, C. H., Elliott, A. & Hanby, W. E. (1956). *Synthetic Polypeptides.* New York: Academic Press.
Blout, E. R. & Fields, M. (1948). *Science*, **107**, 252.
Blout, E. R. & Fields, M. (1949). *J. Biol. Chem.* **173**, 335.
Blout, E. R. & Lenormant, H. (1953). *J. Amer. Chem. Soc.* **43**, 1093.
Bradbury, E. M. (1959). Ph.D. Thesis, London University.
Bradbury, E. M., Burge, R. E., Randall, J. T. & Wilkinson, G. R. (1958). *Disc. Faraday Soc.* No. 25.
Crick, F. H. C. & Watson, J. D. (1954), *Proc. Roy. Soc.* A. **223**, 80.
Elliott, A., Ambrose, E. J. & Temple R. B. (1948). *J. Opt. Soc. Amer.* **35**, 212.

Feughelman, M., Langridge, R., Seeds, W. E., Stokes, A. R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., Barclay, R. K. & Hamilton, L. D. (1955). *Nature*, **175**, 834.
Ford, M. A. (1956). Ph.D. Thesis, London University.
Ford, M. A., Price, W. C., Seeds, W. E. & Wilkinson, G. R. (1958). *J. Opt. Soc. Amer.* **48**, 249.
Franklin, R. E. & Gosling, R. G. (1953). *Acta Cryst.* **6**, 673.
Fraser, R. D. B. (1951). Ph.D. Thesis, London University.
Fraser, M. J. & Fraser, R. D. B. (1951). *Nature*, **167**, 761.
Furberg, S. (1949). Ph.D. Thesis, London University.
Langridge, R., Marvin, D. A., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1960*a*). *J. Mol. Biol.* **2**, 38.
Langridge, R., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1957). *J. Biophys. Biochem.* **3**, 767.
Langridge, R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1960*b*). *J. Mol. Biol.* **2**, 19.
Lenormant, H. & de Lozé, C. (1955). *C.R. Soc. Biol. Paris*, **240**, 1327.
Marvin, D. A., Spencer, M., Wilkins, M. H. F. & Hamilton, L. D. (1958). *Nature*, **182**, 387.
Marvin, D. A., Spencer, M., Wilkins, M. H. F. & Hamilton, L. D. (1961). *J. Mol. Biol.* in the press.
O'Brien, F. E. M. (1948). *J. Sci. Instrum.* **25**, 73.
Seeds, W. E. (1951). Ph.D. Thesis, London University.
Sutherland, G. B. B. M. (1955). *R.C. Ist. Lombardo, Milan*, **89**, 67.
Sutherland, G. B. B. M. & Tsuboi, M. (1957). *Proc. Roy. Soc.* A. **239**, 446.
Tsuboi, M. (1957). *J. Amer. Chem. Soc.* **79**, 1351.
Watson, J. D. & Crick, F. H. C. (1953). *Nature*, **171**, 964.
Wilkins, M. H. F., Stokes, A. R., Wilson, H. R. & Seeds, W. E. (1953) *Nature*, **172**, 759.
Wilkinson, G. R., Price, W. C. & Bradbury, E. M. (1959). *Spectrochimica Acta*, **14**, 284.

J. Mol. Biol. (1961) **3**, 318–356

Review Article

Genetic Regulatory Mechanisms in the Synthesis of Proteins †

François Jacob and Jacques Monod

Services de Génétique Microbienne et de Biochimie Cellulaire, Institut Pasteur, Paris

(*Received 28 December 1960*)

The synthesis of enzymes in bacteria follows a double genetic control. The so-called structural genes determine the molecular organization of the proteins. Other, functionally specialized, genetic determinants, called regulator and operator genes, control the rate of protein synthesis through the intermediacy of cytoplasmic components or repressors. The repressors can be either inactivated (induction) or activated (repression) by certain specific metabolites. This system of regulation appears to operate directly at the level of the synthesis by the gene of a short-lived intermediate, or messenger, which becomes associated with the ribosomes where protein synthesis takes place.

1. Introduction

According to its most widely accepted modern connotation, the word "gene" designates a DNA molecule whose specific self-replicating structure can, through mechanisms unknown, become translated into the specific structure of a polypeptide chain.

This concept of the "structural gene" accounts for the multiplicity, specificity and genetic stability of protein structures, and it implies that such structures are not controlled by environmental conditions or agents. It has been known for a long time, however, that the synthesis of individual proteins may be provoked or suppressed within a cell, under the influence of specific external agents, and more generally that the relative rates at which different proteins are synthesized may be profoundly altered, depending on external conditions. Moreover, it is evident from the study of many such effects that their operation is absolutely essential to the survival of the cell.

It has been suggested in the past that these effects might result from, and testify to, complementary contributions of genes on the one hand, and some chemical factors on the other in determining the final structure of proteins. This view, which contradicts at least partially the " structural gene" hypothesis, has found as yet no experimental support, and in the present paper we shall have occasion to consider briefly some of this negative evidence. Taking, at least provisionally, the structural gene hypothesis in its strictest form, let us assume that the DNA message contained within a gene is both necessary and sufficient to define the structure of a protein. The elective effects of agents other than the structural gene itself in promoting or suppressing the synthesis of a protein must then be described as operations which control the rate of transfer of structural information from gene to protein. Since it seems to be established

† This work has been aided by grants from the National Science Foundation, the Jane Coffin Childs Memorial Fund for Medical Research, and the Commissariat à l'Energie Atomique.

that proteins are synthesized in the cytoplasm, rather than directly at the genetic level, this transfer of structural information must involve a chemical intermediate synthesized by the genes. This hypothetical intermediate we shall call the structural messenger. The rate of information transfer, i.e. of protein synthesis, may then depend either upon the activity of the gene in synthesizing the messenger, or upon the activity of the messenger in synthesizing the protein. This simple picture helps to state the two problems with which we shall be concerned in the present paper. If a given agent specifically alters, positively or negatively, the rate of synthesis of a protein, we must ask:

(a) Whether the agent acts at the cytoplasmic level, by controlling the activity of the messenger, or at the genetic level, by controlling the synthesis of the messenger.

(b) Whether the specificity of the effect depends upon some feature of the information transferred from structural gene to protein, or upon some specialized controlling element, not represented in the structure of the protein, gene or messenger.

The first question is easy to state, if difficult to answer. The second may not appear so straightforward. It may be stated in a more general way, by asking whether the genome is composed exclusively of structural genes, or whether it also involves determinants which may control the rates of synthesis of proteins according to a given set of conditions, without determining the structure of any individual protein. Again it may not be evident that these two statements are equivalent. We hope to make their meaning clear and to show that they are indeed equivalent, when we consider experimental examples.

The best defined systems wherein the synthesis of a protein is seen to be controlled by specific agents are examples of enzymatic adaptation, this term being taken here to cover both enzyme induction, i.e. the formation of enzyme electively provoked by a substrate, and enzyme repression, i.e. the specific inhibition of enzyme formation brought about by a metabolite. Only a few inducible and repressible systems have been identified both biochemically and genetically to an extent which allows discussion of the questions in which we are interested here. In attempting to generalize, we will have to extrapolate from these few systems. Such generalization is greatly encouraged, however, by the fact that lysogenic systems, where phage protein synthesis might be presumed to obey entirely different rules, turn out to be analysable in closely similar terms. We shall therefore consider in succession certain inducible and repressible enzyme systems and lysogenic systems.

It might be best to state at the outset some of the main conclusions which we shall arrive at. These are:

(a) That the mechanisms of control in all these systems are negative, in the sense that they operate by inhibition rather than activation of protein synthesis.

(b) That in addition to the classical structural genes, these systems involve two other types of genetic determinants (regulator and operator) fulfilling specific functions in the control mechanisms.

(c) That the control mechanisms operate at the genetic level, i.e. by regulating the activity of structural genes.

2. Inducible and Repressible Enzyme Systems

(a) *The phenomenon of enzyme induction. General remarks*

It has been known for over 60 years (Duclaux, 1899; Dienert, 1900; Went, 1901) that certain enzymes of micro-organisms are formed only in the presence of their

specific substrate. This effect, later named "enzymatic adaptation" by Karstrom (1938), has been the subject of a great deal of experimentation and speculation. For a long time, "enzymatic adaptation" was not clearly distinguished from the selection of spontaneous variants in growing populations, or it was suggested that enzymatic adaptation and selection represented *alternative* mechanisms for the acquisition of a "new" enzymatic property. Not until 1946 were adaptive enzyme systems shown to be controlled in bacteria by discrete, specific, stable, i.e. genetic, determinants (Monod & Audureau, 1946). A large number of inducible systems has been discovered and studied in bacteria. In fact, enzymes which attack exogeneous substrates are, as a general rule, inducible in these organisms. The phenomenon is far more difficult to study in tissues or cells of higher organisms, but its existence has been established quite clearly in many instances. Very often, if not again as a rule, the presence of a substrate induces the formation not of a single but of several enzymes, sequentially involved in its metabolism (Stanier, 1951).

Most of the fundamental characteristics of the induction effect have been established in the study of the "lactose" system of *Escherichia coli* (Monod & Cohn, 1952; Cohn, 1957; Monod, 1959) and may be summarized in a brief discussion of this system from the biochemical and physiological point of view. We shall return later to the genetic analysis of this system.

(b) *The lactose system of* Escherichia coli

Lactose and other β-galactosides are metabolized in *E. coli* (and certain other enteric bacteria) by the hydrolytic transglucosylase β-galactosidase. This enzyme was isolated from *E. coli* and later crystallized. Its specificity, activation by ions and transglucosylase *vs* hydrolase activity have been studied in great detail (*cf.* Cohn, 1957). We need only mention the properties that are significant for the present discussion. The enzyme is active exclusively on β-galactosides unsubstituted on the galactose ring. Activity and affinity are influenced by the nature of the aglycone moiety both being maximum when this radical is a relatively large, hydrophobic group. Substitution of sulfur for oxygen in the galactosidic linkage of the substrate abolishes hydrolytic activity completely, but the thiogalactosides retain about the same affinity for the enzyme site as the homologous oxygen compounds.

As isolated by present methods, β-galactosidase appears to form various polymers (mostly hexamers) of a fundamental unit with a molecular weight of 135,000. There is one end group (threonine) and also one enzyme site (as determined by equilibrium dialysis against thiogalactosides) per unit. It is uncertain whether the monomer is active as such, or exists *in vivo*. The hexameric molecule has a turnover number of 240,000 mol $\times$ min^{-1} at 28°C, pH 7·0 with *o*-nitrophenyl-β-D-galactoside as substrate and Na^+ (0·01 M) as activator.

There seems to exist only a single homogeneous β-galactosidase in *E. coli*, and this organism apparently cannot form any other enzyme capable of metabolizing lactose, as indicated by the fact that mutants that have lost β-galactosidase activity cannot grow on lactose as sole carbon source.

However, the possession of β-galactosidase activity is not sufficient to allow utilization of lactose by *intact E. coli* cells. Another component, distinct from β-galactosidase, is required to allow penetration of the substrate into the cell (Monod, 1956; Rickenberg, Cohen, Buttin & Monod, 1956; Cohen & Monod, 1957; Pardee, 1957; Képès, 1960). The presence and activity of this component is determined by measuring the rate of

entry and/or the level of accumulation of radioactive thiogalactosides into intact cells. Analysis of this active permeation process shows that it obeys classical enzyme kinetics allowing determination of K_m and V_{max}. The specificity is high since the system is active only with galactosides (β or α), or thiogalactosides. The spectrum of apparent affinities ($1/K_m$) is very different from that of β-galactosidase. Since the permeation system, like β-galactosidase, is inducible (see below) its formation can be studied *in vivo*, and shown to be invariably associated with protein synthesis. By these criteria, there appears to be little doubt that this specific permeation system involves a specific protein (or proteins), formed upon induction, which has been called galactoside-permease. That this protein is distinct from and independent of β-galactosidase is shown by the fact that mutants that have lost β-galactosidase retain the capacity to concentrate galactosides, while mutants that have lost this capacity retain the power to synthesize galactosidase. The latter mutants (called cryptic) cannot however use lactose, since the intracellular galactosidase is apparently accessible exclusively *via* the specific permeation system.

Until quite recently, it had not proved possible to identify *in vitro* the inducible protein (or proteins) presumably responsible for galactoside-permease activity. During the past year, a protein characterized by the ability to carry out the reaction:

Ac. Coenzyme A + Thiogalactoside → 6-Acetylthiogalactoside + Coenzyme A

has been identified, and extensively purified from extracts of *E. coli* grown in presence of galactosides (Zabin, Képès & Monod, 1959). The function of this enzyme in the system is far from clear, since formation of a free covalent acetyl-compound is almost certainly not involved in the permeation process *in vivo*. On the other hand:

(a) mutants that have lost β-galactosidase and retained galactoside-permease, retain galactoside-acetylase;

(b) most mutants that have lost permease cannot form acetylase;

(c) permeaseless acetylaseless mutants which revert to the permease-positive condition simultaneously regain the ability to form acetylase.

These correlations strongly suggest that galactoside-acetylase is somehow involved in the permeation process, although its function *in vivo* is obscure, and it seems almost certain that other proteins (specific or not for this system) are involved. In any case, we are interested here not in the mechanisms of permeation, but in the control mechanisms which operate with β-galactosidase, galactoside-permease and galactoside-acetylase. The important point therefore is that, as we shall see, galactoside-acetylase invariably obeys the same controls as galactosidase.†

(c) *Enzyme induction and protein synthesis*

Wild type *E. coli* cells grown in the absence of a galactoside contain about 1 to 10 units of galactosidase per mg dry weight, that is, an average of 0·5 to 5 active molecules

† For reasons which will become apparent later it is important to consider whether there is any justification for the assumption that galactosidase and acetylase activities might be associated with the same fundamental protein unit. We should therefore point to the following observations:

(a) There are mutants which form galactosidase and no acetylase, and *vice versa*.

(b) Purified acetylase is devoid of any detectable galactosidase activity.

(c) The specificity of the two enzymes is very different.

(d) The two enzymes are easily and completely separated by fractional precipitation.

(e) Acetylase is highly heat-resistant, under conditions where galactosidase is very labile.

(f) Anti-galactosidase serum does not precipitate acetylase; nor does anti-acetylase serum precipitate galactosidase.

There is therefore no ground for the contention that galactosidase and acetylase activities are associated with the same protein.

per cell or 0·15 to 1·5 molecules per nucleus. Bacteria grown in the presence of a suitable inducer contain an average of 10,000 units per mg dry weight. This is the induction effect.

A primary problem, to which much experimental work has been devoted, is whether this considerable increase in specific activity corresponds to the synthesis of entirely "new" enzyme molecules, or to the activation or conversion of pre-existing protein precursors. It has been established by a combination of immunological and isotopic methods that the enzyme formed upon induction:

(a) is distinct, as an antigen, from all the proteins present in uninduced cells (Cohn & Torriani, 1952);

(b) does not derive any significant fraction of its sulfur (Monod & Cohn, 1953; Hogness, Cohn & Monod, 1955) or carbon (Rotman & Spiegelman, 1954) from pre-existing proteins.

The inducer, therefore, brings about the complete *de novo* synthesis of enzyme molecules which are new by their specific structure as well as by the origin of their elements. The study of several other induced systems has fully confirmed this conclusion, which may by now be considered as part of the *definition* of the effect. We will use the term "induction" here as meaning "activation by inducer of enzyme-protein synthesis."

(d) *Kinetics of induction*

Accepting (still provisionally) the structural gene hypothesis, we may therefore consider that the inducer somehow accelerates the rate of information transfer from gene to protein. This it could do either by provoking the synthesis of the messenger or by activating the messenger. If the messenger were a *stable* structure, functioning as a catalytic template in protein synthesis, one would expect different kinetics of induction, depending on whether the inducer acted at the genetic or at the cytoplasmic level.

The kinetics of galactosidase induction turn out to be remarkably simple when determined under proper experimental conditions (Monod, Pappenheimer & Cohen-Bazire, 1952; Herzenberg, 1959). Upon addition of a suitable inducer to a growing culture, enzyme activity increases at a rate proportional to the increase in total protein within the culture; i.e. a linear relation is obtained (Fig. 1) when total enzyme activity is plotted against mass of the culture. The slope of this line:

$$P = \frac{\Delta z}{\Delta M}$$

is the "differential rate of synthesis," which is taken by definition as the measure of the effect. Extrapolation to the origin indicates that enzyme formation begins about three minutes (at 37°C) after addition of inducer (Pardee & Prestidge, 1961). Removal of the inducer (or addition of a specific anti-inducer, see below) results in cessation of enzyme synthesis within the same short time. The differential rate of synthesis varies with the concentration of inducer reaching a different saturation value for different inducers. The inducer therefore acts in a manner which is (kinetically) similar to that of a dissociable activator in an enzyme system: activation and inactivation follow very rapidly upon addition or removal of the activator.

The conclusion which can be drawn from these kinetics is a negative one: the inducer does not appear to activate the synthesis of a stable intermediate able to accumulate in the cell (Monod, 1956).

Similar kinetics of induction have been observed with most or all other systems which have been adequately studied (Halvorson, 1960) with the exception of penicillinase of *Bacillus cereus*. The well-known work of Pollock has shown that the synthesis of this enzyme continues for a long time, at a decreasing rate, after removal of inducer (penicillin) from the medium. This effect is apparently related to the fact that minute amounts of penicillin are retained irreversibly by the cells after transient exposure to the drug (Pollock, 1950). The unique behavior of this system therefore does not contradict the rule that induced synthesis stops when the inducer is removed from the cells. Using this system, Pollock & Perret (1951) were able to show that the inducer acts catalytically, in the sense that a cell may synthesize many more enzyme molecules than it has retained inducer molecules.

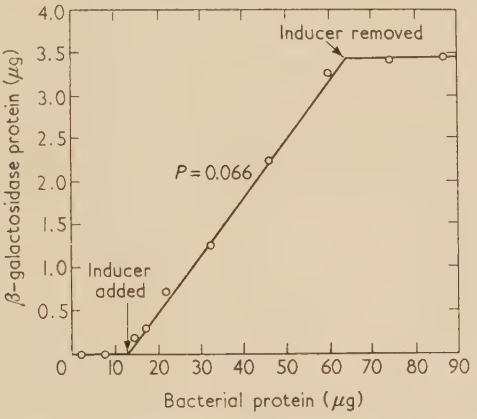


FIG. 1. Kinetics of induced enzyme synthesis. Differential plot expressing accumulation of β-galactosidase as a function of increase of mass of cells in a growing culture of *E. coli*. Since abscissa and ordinates are expressed in the same units (micrograms of protein) the slope of the straight line gives galactosidase as the fraction (P) of total protein synthesized in the presence of inducer. (After Cohn, 1957.)

(e) *Specificity of induction*

One of the most conspicuous features of the induction effect is its extreme specificity. As a general rule, only the substrate of an enzyme, or substances very closely allied to the normal substrate, are endowed with inducer activity towards this enzyme. This evidently suggests that a correlation between the molecular structure of the inducer and the structure of the catalytic center on the enzymes is *inherently* involved in the mechanism of induction. Two main types of hypotheses have been proposed to account for this correlation, and thereby for the mechanism of action of the inducer:

(a) The inducer serves as "partial template" in enzyme synthesis, molding as it were the catalytic center.

(b) The inducer acts by combining specifically with preformed enzyme (or "pre-enzyme"), thereby somehow accelerating the synthesis of further enzyme molecules.

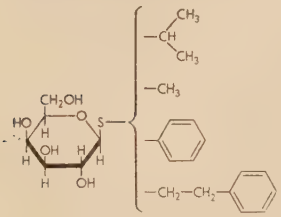
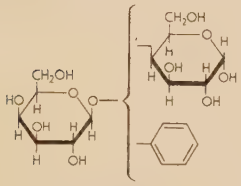
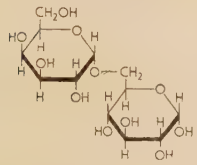
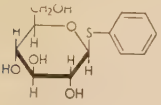
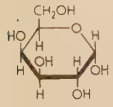
It is not necessary to discuss these "classical" hypotheses in detail, because it seems to be established now that the correlation in question is in fact *not* inherent to the mechanism of induction.

Table 1 lists a number of compounds tested as inducers of galactosidase, and as substrates (or specific inhibitors) of the enzyme. It will be noted that:

(a) no compound that does not possess an intact unsubstituted galactosidic residue induces;

TABLE 1

Induction of galactosidase and galactoside-transacetylase by various galactosides

Compound	Concentrations	β-galactosidase			Galactoside-transacetylase	
		Induction value	V	$1/K_m$	Induction value	V/K_m
β-D-thiogalactosides						
 (*iso*propyl)	10^{-4} M	100	0	140	100	80
(methyl)	10^{-4} M	78	0	7	74	30
	10^{-5} M	7·5	—	—	10	—
(phenyl)	10^{-3} M	<0·1	0	100	<1	100
(phenylethyl)	10^{-3} M	5	0	10,000	3	—
β-D-galactosides						
 (lactose)	10^{-3} M	17	30	14	12	35
(phenyl)	10^{-3} M	15	100	100	11	—
α-D-galactoside						
 (melibiose)	10^{-3} M	35	0	<0·1	37	<1
β-D-glucoside						
 (phenyl)	10^{-3} M	<0·1	0	0	<1	50
 (galactose)	10^{-3} M	<0·1	—	4	<1	<1
Methyl-β-D-thiogalactoside (10^{-4} M) + phenyl-β-D-thiogalactoside (10^{-3} M)		52	—	—	63	—

Columns "induction value" refer to specific activities developed by cultures of wild type *E. coli* K12 grown on glycerol as carbon source with each galactoside added at molar concentration stated. Values are given in percent of values obtained with *iso*propyl-thiogalactoside at 10^{-4} M (for which actual units were about 7,500 units of β-galactosidase and 300 units of galactoside-transacetylase per mg of bacteria). Column V refers to maximal substrate activity of each compound with respect to galactosidase. Values are given in percent of activity obtained with phenyl-galactoside. Column $1/K_m$ expresses affinity of each compound with respect to galactosidase. Values are given in percent of that observed with phenylgalactoside. In case of galactoside-transacetylase, only the relative values V/K_m are given since low affinity of this enzyme prevents independent determination of the constants. (Computed from Monod & Cohn, 1952; Monod *et al.*, 1952; Buttin, 1956; Zabin *et al.*, 1959; Képès *et al.*, unpublished results.)

(b) many compounds which are not substrates (such as the thiogalactosides) are excellent inducers (for instance *iso*propyl thiogalactoside);

(c) there is no correlation between affinity for the enzyme and capacity to induce (*cf.* thiophenylgalactoside and melibiose).

The possibility that the enzyme formed in response to different inducers may have somewhat different specific properties should also be considered, and has been rather thoroughly tested, with entirely negative results (Monod & Cohn, 1952).

There is therefore no quantitative correlation whatever between inducing capacity and the substrate activity or affinity parameters of the various galactosides tested. The fact remains, however, that only galactosides will induce galactosidase, whose binding site is complementary for the galactose ring-structure. The possibility that this correlation is a necessary requisite, or consequence, of the induction mechanism was therefore not completely excluded by the former results.

As we shall see later, certain mutants of the galactosidase structural gene (z) have been found to synthesize, in place of the normal enzyme, a protein which is identical to it by its immunological properties, while being completely devoid of any enzymatic activity. When tested by equilibrium dialysis, this inactive protein proved to have no measurable affinity for galactosides. In other words, it has lost the specific binding site. In diploids carrying both the normal and the mutated gene, both normal galactosidase and the inactive protein are formed, to a quantitatively similar extent, in the presence of different concentrations of inducer (Perrin, Jacob & Monod, 1960).

This finding, added to the sum of the preceding observations, appears to prove beyond reasonable doubt that the mechanism of induction does not imply any inherent correlation between the molecular structure of the inducer and the structure of the binding site of the enzyme.

On the other hand, there is complete correlation in the induction of galactosidase and acetylase. This is illustrated by Table 1 which shows not only that the same compounds are active or inactive as inducers of either enzyme, but that the relative amounts of galactosidase and acetylase synthesized in the presence of different inducers or at different concentrations of the same inducer are constant, even though the absolute amounts vary greatly. The remarkable qualitative and quantitative correlation in the induction of these two widely different enzyme proteins strongly suggests that the synthesis of both is directly governed by a common controlling element with which the inducer interacts. This interaction must, at some point, involve stereospecific binding of the inducer, since induction is sterically specific, and since certain galactosides which are devoid of any inducing activity act as competitive inhibitors of induction in the presence of active inducers (Monod, 1956; Herzenberg, 1959). This suggests that an enzyme, or some other protein, distinct from either galactosidase or acetylase, acts as "receptor" of the inducer. We shall return later to the difficult problem raised by the identification of this "induction receptor."

(f) *Enzyme repression*

While positive enzymatic adaptation, i.e. induction, has been known for over sixty years, negative adaptation, i.e. specific inhibition of enzyme synthesis, was discovered only in 1953, when it was found that the formation of the enzyme tryptophan-synthetase was inhibited selectively by tryptophan and certain tryptophan analogs (Monod & Cohen-Bazire, 1953). Soon afterwards, other examples of this effect were observed (Cohn, Cohen & Monod, 1953; Adelberg & Umbarger, 1953; Wijesundera &

Woods, 1953), and several systems were studied in detail in subsequent years (Gorini & Maas, 1957; Vogel, 1957*a*,*b*; Yates & Pardee, 1957; Magasanik, Magasanik & Neidhardt, 1959). These studies have revealed that the "repression" effect, as it was later named by Vogel (1957*a*,*b*), is very closely analogous, albeit symmetrically opposed, to the induction effect.

Enzyme repression, like induction, generally involves not a single but a sequence of enzymes active in successive metabolic steps. While inducibility is the rule for catabolic enzyme sequences responsible for the degradation of exogeneous substances, repressibility is the rule for anabolic enzymes, involved in the synthesis of essential metabolites such as amino acids or nucleotides.† Repression, like induction, is highly specific, but while inducers generally are substrates (or analogs of substrates) of the sequence, the repressing metabolites generally are the product (or analogs of the product) of the sequence.

That the effect involves inhibition of enzyme *synthesis*, and not inhibition (directly or indirectly) of enzyme *activity* was apparent already in the first example studied (Monod & Cohen-Bazire, 1953), and has been proved conclusively by isotope incorporation experiments (Yates & Pardee, 1957). It is important to emphasize this point, because enzyme repression must not be confused with another effect variously called "feedback inhibition" or "retro-inhibition" which is equally frequent, and may occur in the same systems. This last effect, discovered by Novick & Szilard (in Novick, 1955), involves the inhibition of activity of an early enzyme in an anabolic sequence, by the ultimate product of the sequence (Yates & Pardee, 1956; Umbarger, 1956). We shall use "repression" exclusively to designate specific inhibition of enzyme *synthesis*.‡

(g) *Kinetics and specificity of repression*

The kinetics of enzyme synthesis provoked by "de-repression" are identical to the kinetics of induction (see Fig. 2). When wild type *E. coli* is grown in the presence of arginine, only traces of ornithine-carbamyltransferase are formed. As soon as arginine is removed from the growth medium, the differential rate of enzyme synthesis increases about 1,000 times and remains constant, until arginine is added again, when it immediately falls back to the repressed level. The repressing metabolite here acts (kinetically) as would a dissociable inhibitor in an enzyme system.

The specificity of repression poses some particularly significant problems. As a rule, the repressing metabolite of an anabolic sequence is the ultimate product of this sequence. For instance, L-arginine, to the exclusion of any other amino acid, represses the enzymes of the sequence involved in the biosynthesis of arginine. Arginine shows no specific affinity for the early enzymes in the sequence, such as, in particular, ornithine-carbamyltransferase. In this sense, arginine is a "gratuitous" repressing metabolite for this protein, just as galactosides are "gratuitous inducers" for the mutated (inactive) galactosidase. The possibility must be considered however that arginine may be converted back, through the sequence itself, to an intermediate product

† Certain enzymes which attack exogeneous substrates are controlled by repression. Alkaline phosphatase (*E. coli*) is not induced by phosphate esters, but it is repressed by orthophosphate. Urease (*Pseudomonas*) is repressed by ammonia.

‡ We should perhaps recall the well-known fact that glucose and other carbohydrates inhibit the synthesis of many *inducible* enzymes, attacking a variety of substrates (Dienert, 1900; Gale, 1943; Monod, 1942; Cohn & Horibata, 1959). It is probable that this non-specific "glucose effect" bears some relation to the repressive effect of specific metabolites, but the relationship is not clear (Neidhardt & Magasanik, 1956*a*,*b*). We shall not discuss the glucose effect in this paper.

or substrate of the enzyme. This has been excluded by Gorini & Maas (1957) who showed that, in mutants lacking one of the enzymes involved in later steps of the sequence, ornithine transcarbamylase is repressed by arginine to the same extent as in the wild type. Moreover, neither ornithine nor any other intermediate of the sequence is endowed with repressing activity in mutants which cannot convert the intermediate into arginine. It is quite clear therefore that the specificity of action of the repressing metabolite does not depend upon the specific configuration of the enzyme site.

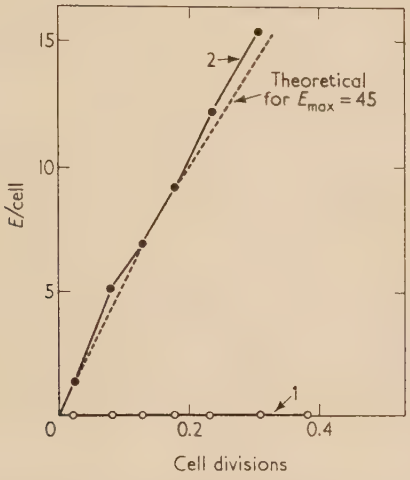


FIG. 2. Repression of ornithine-transcarbamylase by arginine. *E. coli* requiring both histidine and arginine were grown in a chemostat with 1 μg/ml. histidine + 6 μg/ml. arginine (curve 1) or with 10 μg/ml. histidine + 5 μg/ml. arginine (curve 2). Cultures are inoculated with washed cells taken from cultures growing exponentially in excess of arginine. The theoretical curve was calculated from the constant enzyme/cell value reached after 4 cell divisions. (After Gorini & Maas, 1958.)

The same conclusion is applicable to the enzymes of the histidine synthesizing pathway which are repressed in the presence of histidine, both in the wild type and in different mutants lacking one of the enzymes. The work of Ames & Garry (1959) has shown that the rates of synthesis of different enzymes in this sequence vary in *quantitatively* constant ratios under any set of medium conditions, and that the ratios are the same in various mutants lacking one of the enzymes and in the wild type. Here again, as in the case of the lactose system, the synthesis of widely different, albeit functionally related, enzymes appears to be controlled by a single common mechanism, with which the repressing metabolite specifically interacts.

In summary, repression and induction appear as closely similar effects, even if opposed in their results. Both control the rate of synthesis of enzyme proteins. Both are highly specific, but in neither case is the specificity related to the specificity of action (or binding) of the controlled enzyme. The kinetics of induction and repression are the same. Different functionally related enzymes are frequently co-induced or co-repressed, quantitatively to the same extent, by a single substrate or metabolite.

The remarkable similarity of induction and repression suggests that the two effects represent different manifestations of fundamentally similar mechanisms (Cohn & Monod, 1953; Monod, 1955; Vogel, 1957*a*, *b*; Pardee, Jacob & Monod, 1959; Szilard,

1960). This would imply either that in inducible systems the inducer acts as an antagonist of an internal repressor or that in repressible systems the repressing metabolite acts as an antagonist of an internal inducer. This is not an esoteric dilemma since it poses a very pertinent question, namely what would happen in an adaptive system of either type, when *both* the inducer and the repressor were eliminated? This, in fact, is the main question which we shall try to answer in the next section.

3. Regulator Genes

Since the specificity of induction or repression is not related to the structural specificity of the controlled enzymes, and since the rate of synthesis of different enzymes appears to be governed by a common element, this element is presumably not controlled or represented by the structural genes themselves. This inference, as we shall now see, is confirmed by the study of certain mutations which convert inducible or repressible systems into constitutive systems.

(a) *Phenotypes and genotypes in the lactose systems*

If this inference is correct, mutations which affect the controlling system should not behave as alleles of the structural genes. In order to test this prediction, the structural genes themselves must be identified. The most thoroughly investigated case is the lactose system of *E. coli*, to which we shall now return. Six phenotypically different classes of mutants have been observed in this system. For the time being, we shall consider only three of them which will be symbolized and defined as follows:

(1) Galactosidase mutations: $z^+ \rightleftharpoons z^-$ expressed as the loss of the capacity to synthesize active galactosidase (with or without induction).

(2) Permease mutations: $y^+ \rightleftharpoons y^-$ expressed as the loss of the capacity to form galactoside-permease. Most, but not all, mutants of this class simultaneously lose the capacity to synthesize active acetylase. We shall confine our discussion to the acetylaseless subclass.

(3) Constitutive mutations: $i^+ \rightleftharpoons i^-$ expressed as the ability to synthesize large amounts of galactosidase *and* acetylase in the absence of inducer (Monod, 1956; Rickenberg *et al.*, 1956; Pardee *et al.*, 1959).

The first two classes are specific for either galactosidase or acetylase: the galactosidaseless mutants form normal amounts of acetylase; conversely the acetylaseless mutants form normal amounts of galactosidase. In contrast, the constitutive mutations, of which over one hundred recurrences have been observed, invariably affect both the galactosidase and the permease (acetylase).† There are eight possible combinations of these phenotypes, and they have all been observed both in *E. coli* ML and K12.

The loci corresponding to a number of recurrences of each of the three mutant types have been mapped by recombination in *E. coli* K12. The map (Fig. 3) also

† The significance of this finding could be questioned since, in order to isolate constitutive mutants, one must of course use selective media, and this procedure might be supposed to favour double mutants, where the constitutivity of galactosidase and permease had arisen independently. It is possible, however, to select for $i^+ \rightarrow i^-$ mutants in organisms of type $i^+z^+y^-$, i.e. permeaseless. Fifty such mutants were isolated, giving rise to "constitutive cryptic" types $i^-z^+y^-$ from which, by reversion of y^-, fifty clones of constitutive $i^-z^+y^+$ were obtained. It was verified that in each of these fifty clones the permease was constitutive.

indicates the location of certain other mutations (*o* mutations) which will be discussed later. As may be seen, all these loci are confined to a very small segment of the chromosome, the *Lac* region. The extreme proximity of all these mutations raises the question whether they belong to a single or to several independent functional units. Such functional analysis requires that the biochemical expression of the various genetic structures be studied in heterozygous diploids. Until quite recently, only transient diploids were available in *E. coli*; the recent discovery of a new type of gene transfer in these bacteria (sexduction) has opened the possibility of obtaining stable clones which are diploid (or polyploid) for different small segments of the chromosome.

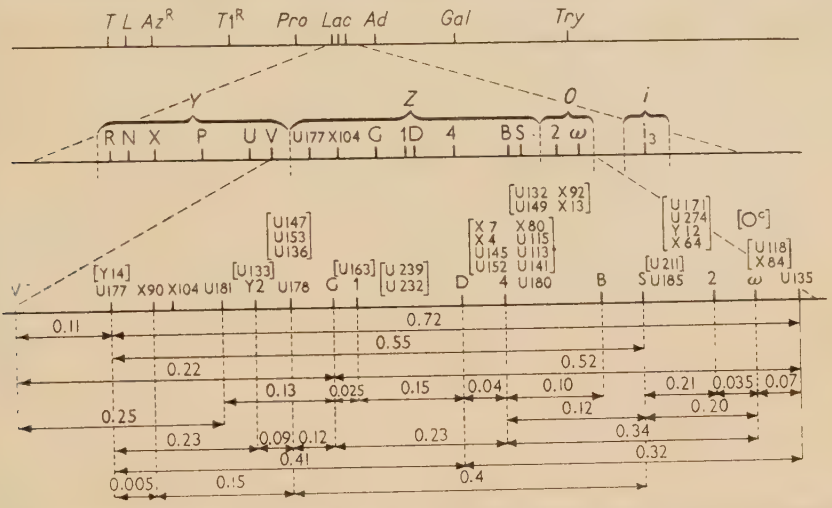


FIG. 3. Diagrammatic map of the lactose region of *E. coli* K12. The upper line represents the position of the *Lac* region with respect to other known markers. The middle line represents an enlargement of the *Lac* region with the four loci *y*, *z*, *o* and *i*. The lower line represents an enlargement of the *z* and *o* loci. Recombination frequencies (given at the bottom) are obtained in two factor crosses of the type $Hfr\ Lac^{-}_{A}ad^{+}S^{s} \times F^{-}Lac^{-}_{B}ad^{-}S^{r}$, from the ratios "recombinants $Lac^{+}ad^{+}S^{r}$/recombinants $ad^{+}S^{r}$." The total length of the *z* gene may be estimated to be 0·7 map units, i.e. about 3,500 nucleotide pairs for about 1,000 amino acids in the monomer of β-galactosidase.

In this process, small fragments of the bacterial chromosome are incorporated into the sex factor, *F*. This new unit of replication is transmissible by conjugation, and is then added to the normal genome of the recipient bacterium which becomes diploid for the small chromosomal fragment. Among the units thus isolated, one carries the whole *Lac* region (Jacob & Adelberg, 1959; Jacob, Perrin, Sanchez & Monod, 1960). To symbolize the genetic structure of these diploids, the chromosomal alleles are written in the usual manner, while the alleles attached to the sex factor are preceded by the letter *F*.

Turning our attention to the behaviour of *z* and *y* mutant types, we may first note that diploids of structure $z^{+}y^{-}/Fz^{-}y^{+}$ or $z^{-}y^{+}/Fz^{+}y^{-}$ are wild type, being able to ferment lactose, and forming normal amounts of both galactosidase and acetylase. This complete complementation between z^{-} and y^{-} mutants indicates that they belong to independent cistrons. Conversely, no complementation is observed between different y^{-} mutants, indicating that they all belong to a single cistron. No complementation is observed between most z^{-} mutants. Certain diploids of structure $z^{-}_{a}z^{+}_{b}/Fz^{+}_{a}z^{-}_{b}$ synthesize galactosidase in reduced amounts, but pairs of mutually non-complementing mutants overlap mutually complementing mutants, suggesting again

that a single cistron is involved, as one might expect, since the monomer of galactosidase has a single N-terminal group. It should be recalled that intracistronic partial complementation has been observed in several cases (Giles, 1958), and has (tentatively) been explained as related to a polymeric state of the protein.

Mutations in the *z* gene affect the structure of galactosidase. This is shown by the fact that most of the z^- mutants synthesize, in place of active enzyme, a protein which is able to displace authentic (wild type) galactosidase from its combination with specific antibody (Perrin, Bussard & Monod, 1959). Among proteins synthesized by different z^- mutants (symbolized Cz_1, Cz_2, etc.) some give complete cross reactions (i.e. precipitate 100% of the specific antigalactosidase antibodies) with the serum used, while others give incomplete reactions. The different Cz proteins differ therefore, not only from wild type galactosidase, but also one from the other. Finally, as we already mentioned, diploids of constitution z^+/z_1^- synthesize wild type galactosidase and the modified protein simultaneously, and at similar rates (Perrin *et al.*, 1960). These observations justify the conclusions that the *z* region or cistron contains the structural information for β-galactosidase. Proof that mutations in the *y* region not only suppress but may in some cases modify the structure of acetylase has not been obtained as yet, but the assumption that the *y* region does represent, in part at least, the structural gene for the acetylase protein appears quite safe in view of the properties of the *y* mutants.

(b) *The* i$^+$ *gene and its cytoplasmic product*

We now turn our attention to the constitutive (i^-) mutations. The most significant feature of these mutations is that they invariably affect simultaneously two different enzyme-proteins, each independently determined, as we have just seen, by different structural genes. In fact, most i^- mutants synthesize more galactosidase and acetylase than induced wild type cells, but it is quite remarkable that the *ratio* of galactosidase to acetylase is the same in the constitutive cells as in the induced wild type, strongly suggesting that the mechanism controlled by the *i* gene is the same as that with which the inducer interacts.

The study of double heterozygotes of structures: i^+z^-/Fi^-z^+ or i^-y^+/Fi^+y^- shows (Table 2, lines 4 and 5) that the inducible i^+ allele is dominant over the constitutive and that it is active in the *trans* position, with respect to both y^+ and z^+.

Therefore the *i* mutations belong to an independent cistron, governing the expression of *y* and *z via* a cytoplasmic component. The dominance of the inducible over the constitutive allele means that the former corresponds to the active form of the *i* gene. This is confirmed by the fact that strains carrying a *deletion* of the *izy* region behave like i^- in diploids (Table 2, line 7). However, two different interpretations of the function of the i^+ gene must be considered.

(a) The i^+ gene determines the synthesis of a repressor, inactive or absent in the i^- alleles.

(b) The i^+ gene determines the synthesis of an enzyme which destroys an inducer, produced by an independent pathway.

The first interpretation is the most straightforward, and it presents the great interest of implying that the fundamental mechanisms of control may be the same in inducible and repressible systems. Several lines of evidence indicate that it is the correct interpretation.

First, we may mention the fact that constitutive synthesis of β-galactosidase by $i^-z^+y^+$ types is not inhibited by thiophenyl-galactoside which has been shown (Cohn & Monod, 1953) to be a competitive inhibitor of induction by exogenous galactosides (see p. 325).

TABLE 2

Synthesis of galactosidase and galactoside-transacetylase by haploids and heterozygous diploids of regulator mutants

Strain No.	Genotype	Galactosidase		Galactoside-transacetylase	
		Non-induced	Induced	Non-induced	Induced
1	$i^+z^+y^+$	<0·1	100	<1	100
2	$i_6^-z^+y^+$	100	100	90	90
3	$i_3^-z^+y^+$	140	130	130	120
4	$i^+z_1^-y^+/Fi_3^-z^+y^+$	<1	240	1	270
5	$i_3^-z_1^-y^+/Fi^+z^+y\bar{U}$	<1	280	<1	120
6	$i_3\,z_1^-y^+/Fi^-z^+y^+$	195	190	200	180
7	$\Delta_{izy}/Fi^-z^+y^+$	130	150	150	170
8	$i^sz^+y^+$	<0·1	<1	<1	<1
9	$i^sz^+y^+/Fi^+z^+y^+$	<0·1	2	<1	3

Bacteria are grown in glycerol as carbon source and induced, when stated, by *iso*propyl-thio-galactoside, 10^{-4} M. Values are given as a percentage of those observed with induced wild type (for absolute values, see legend of Table 1). Δ_{izy} refers to a deletion of the whole *Lac* region. It will be noted that organisms carrying the wild allele of one of the structural genes (*z* or *y*) on the *F* factor form more of the corresponding enzyme than the haploid. This is presumably due to the fact that several copies of the *F-Lac* unit are present per chromosome. In i^+/i^- heterozygotes, values observed with uninduced cells are sometimes higher than in the haploid control. This is due to the presence of a significant fraction of i^-/i^- homozygous recombinants in the population.

A direct and specific argument comes from the study of one particular mutant of the lactose system. This mutant (i^s) has lost the capacity to synthesize *both* galactosidase and permease. It is not a deletion because it recombines, giving Lac^+ types, with all the z^- and y^- mutants. In crosses with z^-i^- organisms the progeny is *exclusively* i^- while in crosses with z^-i^+ it is *exclusively* i^+, indicating exceedingly close linkage of this mutation with the *i* region. Finally, in diploids of constitution i^s/i^+, i^s turns out to be *dominant*: the diploids cannot synthesize either galactosidase or acetylase (see Table 2, lines 8 and 9).

These unique properties appear exceedingly difficult to account for, except by the admittedly very specific hypothesis that mutant i^s is an allele of *i* where the *structure* of the repressor is such that it cannot be antagonized by the inducer any more. If this hypothesis is correct, one would expect that the i^s mutant could regain the ability to metabolize lactose, not only by reversion to wild type ($i^s \rightarrow i^+$) but also, and probably more frequently, by inactivation of the *i* gene, that is to say by achieving the

constitutive condition ($i^s \rightarrow i^-$). Actually, *Lac*$^+$ "revertants" are very frequent in populations of mutant i^s, and 50% of these "revertants" are indeed constitutives of the i^- (recessive) type. (The other revertants are also constitutives, but of the o^c class which we shall mention later.) The properties of this remarkable mutant could evidently not be understood under the assumption that the *i* gene governs the synthesis of an inducer-destroying enzyme (Willson, Perrin, Jacob & Monod, 1961).

Accepting tentatively the conclusion that the i^+ gene governs the synthesis of an intracellular repressor, we may now consider the question of the presence of this substance in the cytoplasm, and of its chemical nature.

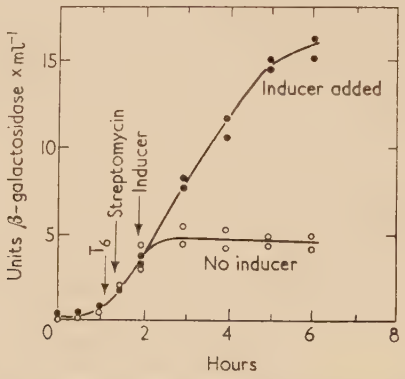


FIG. 4. Synthesis of β-galactosidase by merozygotes formed by conjugation between inducible, galactosidase-positive males and constitutive, galactosidase-negative females. Male (*Hfr* $i^+z^+T6^sS^s$) and female (F^- $i^-z^-T6^rS^r$) bacteria grown in a synthetic medium containing glycerol as carbon source are mixed in the same medium (time O) in the absence of inducer. In such a cross, the first zygotes which receive the *Lac* region from the males are formed from the 20th min. The rate of enzyme synthesis is determined from enzyme activity measurement on the whole population, to which streptomycin and phage T6 are added at times indicated by arrows to block further formation of recombinants and induction of the male parents. It may be seen that in the absence of inducer enzyme synthesis stops about 60 to 80 min after penetration of the first z^+i^+ segment but is resumed by addition of inducer (From Pardee *et al*, 1959).

Important indications on this question have been obtained by studying the kinetics and conditions of expression of the i^+ and z^+ genes when they are introduced into the cytoplasm of cells bearing the inactive (z^- and i^-) alleles. The sexual transfer of the *Lac* segment from male to female cells provides an adequate experimental system for such studies. It should be recalled that conjugation in *E. coli* involves essentially the transfer of a male chromosome (or chromosome segment) to the female cell. This transfer is oriented, always beginning at one extremity of the chromosome, and it is progressive, each chromosome segment entering into the recipient cell at a fairly precise time following inception of conjugation in a given mating pair (Wollman & Jacob, 1959). The conjugation does not appear to involve any significant cytoplasmic mixing, so that the zygotes inherit virtually all their cytoplasm from the female cell, receiving only a chromosome or chromosome segment from the male. In order to study galactosidase synthesis by the zygotes, conditions must be set up such that the unmated parents cannot form the enzymes. This is the case when mating between inducible galactosidase-positive, streptomycin-sensitive males (δ $z^+i^+Sm^s$) and constitutive, galactosidase-negative, streptomycin-resistant females (φ $z^-i^-Sm^r$) is performed in presence of streptomycin (Sm), since: (i) the male cells which are sensitive to Sm cannot synthesize enzyme in its presence; (ii) the female cells are genetically incompetent; (iii) the vast majority of the zygotes which receive the z^+ gene, do not

become streptomycin sensitive (because the Sm^s gene is transferred only to a small proportion of them, and at a very late time). The results of such an experiment, performed in the absence of inducer, are shown in Fig. 4. It is seen that galactosidase synthesis starts almost immediately following actual entry of the z^+ gene. We shall return later to a more precise analysis of the expression of the z^+ gene. The important point to be stressed here is that during this initial period the zygotes behave like *constitutive* cells, synthesizing enzyme in the *absence* of inducer. Approximately sixty minutes later, however, the rate of galactosidase synthesis falls off to zero. If at that time inducer is added, the maximum rate of enzyme synthesis is resumed. We are, in other words, witnessing the conversion of the originally i^- phenotype of the zygote cell, into an i^+ phenotype. And this experiment clearly shows that the "inducible" state is associated with the presence, at a sufficient level, of a *cytoplasmic* substance synthesized under the control of the i^+ gene. (It may be pointed out that the use of a female strain carrying a *deletion* of the *Lac* region instead of the i^-z^- alleles gives the same results (Pardee *et al.*, 1959).)

If now 5-methyltryptophan is added to the mated cells a few minutes before entry of the z^+ gene, no galactosidase is formed because, as is well known, this compound inhibits tryptophan synthesis by retro-inhibition, and therefore blocks protein synthesis. If the repressor is a protein, or if it is formed by a specific enzyme, the synthesis of which is governed by the i^+ gene, its accumulation should also be blocked. If on the other hand the repressor is not a protein, and if its synthesis does not require the preliminary synthesis of a specific enzyme controlled by the i^+ gene, it may accumulate in presence of 5-methyltryptophan which is known (Gros, unpublished results) *not* to inhibit energy transfer or the synthesis of nucleic acids.

The results of Pardee & Prestidge (1959) show that the repressor *does* accumulate under these conditions, since the addition of tryptophan 60 min after 5-methyltryptophan allows immediate and complete resumption of enzyme synthesis, *but only in the presence of inducer*; in other words, the cytoplasm of the zygote cells has been converted from the constitutive to the inducible state during the time that protein synthesis was blocked. This result has also been obtained using chloramphenicol as the agent for blocking protein synthesis, and it has been repeated using another system of gene transfer (Luria *et al.*, unpublished results).

This experiment leads to the conclusion that the repressor is not a protein, and this again excludes the hypothesis that the i^+ gene controls an inducer-destroying enzyme. We should like to stress the point that this conclusion does not imply that no enzyme is involved in the synthesis of the repressor, but that the enzymes which may be involved are *not* controlled by the i^+ gene. The experiments are negative, as far as the chemical nature of the repressor itself is concerned, since they only eliminate protein as a candidate. They do, however, invite the speculation that the repressor may be the primary product of the i^+ gene, and the further speculation that such a primary product may be a polyribonucleotide.

Before concluding this section, it should be pointed out that constitutive mutations have been found in several inducible systems; in fact wherever they have been searched for by adequate selective techniques (amylomaltase of *E. coli* (Cohen-Bazire & Jolit, 1953), penicillinase of *B. cereus* (Kogut, Pollock & Tridgell, 1956), glucuronidase of *E. coli* (F. Stoeber, unpublished results), galactokinase and galactose-transferase (Buttin, unpublished results)). That *any* inducible system should be potentially capable of giving rise to constitutive mutants, strongly indicates that such mutations occur, or at least can always occur, by a loss of function. In the case of the "galactose"

system of *E. coli*, it has been found that the constitutive mutation is pleiotropic, affecting a sequence of three different enzymes (galactokinase, galactose-transferase, UDP-galactose epimerase), and occurs at a locus distinct from that of the corresponding structural genes (Buttin, unpublished results).

The main conclusions from the observations reviewed in this section may be summarized as defining a new type of gene, which we shall call a "regulator gene" (Jacob & Monod, 1959). A regulator gene does not contribute structural information to the proteins which it controls. The specific product of a regulator gene is a cytoplasmic substance, which inhibits information transfer from a structural gene (or genes) to protein. In contrast to the classical structural gene, a regulator gene may control the synthesis of several different proteins: the one-gene one-protein rule does not apply to it.

We have already pointed out the profound similarities between induction and repression which suggest that the two effects represent different manifestations of the same fundamental mechanism. If this is true, and if the above conclusions are valid, one expects to find that the genetic control of repressible systems also involves regulator genes.

(c) *Regulator genes in repressible systems*

The identification of constitutive or "de-repressed" mutants of several repressible systems has fulfilled this expectation. For the selection of such mutants, certain analogs of the normal repressing metabolite may be used as specific selective agents, because they cannot substitute for the metabolite, except as repressing metabolites. For instance, 5-methyltryptophan does not substitute for tryptophan in protein synthesis (Munier, unpublished results), but it represses the enzymes of the tryptophan-synthesizing sequence (Monod & Cohen-Bazire, 1953). Normal wild type *E. coli* does not grow in the presence of 5-methyltryptophan. Fully resistant stable mutants arise, however, a large fraction of which turn out to be constitutive for the tryptophan system.† The properties of these organisms indicate that they arise by mutation of a regulator gene R_T (Cohen & Jacob, 1959). In these mutants tryptophan-synthetase as well as at least two of the enzymes involved in previous steps in the sequence are formed at the same rate irrespective of the presence of tryptophan, while in the wild type all these enzymes are strongly repressed. Actually the mutants form more of the enzymes in the presence of tryptophan, than does the wild type in its absence (just as i^-z^+ mutants form more galactosidase in the absence of inducer than the wild type does at saturating concentration of inducer). The capacity of the mutants to concentrate tryptophan from the medium is not impaired, nor is their tryptophanase activity increased. The loss of sensitivity to tryptophan as repressing metabolite cannot therefore be attributed to its destruction by, or exclusion from, the cells, and can only reflect the breakdown of the control system itself. Several recurrences of the R_T mutation have been mapped. They are all located in the same small section of the chromosome, at a large distance from the cluster of genes which was shown by Yanofsky & Lennox (1959) to synthesize the different enzymes of the sequence. One of these genes (comprising two cistrons) has been very clearly identified by the work of Yanofsky (1960) as the structural gene for tryptophan synthetase, and it is a safe assumption that the other genes in this cluster determine the structure of the preceding

† Resistance to 5-methyltryptophan may also arise by other mechanisms in which we are not interested here.

enzymes in the sequence. The R_T gene therefore controls the rate of synthesis of several different proteins without, however, determining their structure. It can only do so *via* a cytoplasmic intermediate, since it is located quite far from the structural genes. To complete its characterization as a regulator gene, it should be verified that the constitutive (R_T^-) allele corresponds to the inactive state of the gene (or gene product), i.e. is recessive. Stable heterozygotes have not been available in this case, but the transient (sexual) heterozygotes of a cross $\sigma^{\!\!\!

to account for such antagonistic or activating interactions. We shall not go into these speculations since there is at present no evidence to support or eliminate any particular model. But it must be pointed out that, in any model, the structural specificity of inducers or co-repressors must be accounted for, and can be accounted for, only by the assumption that a stereospecific receptor is involved in the interaction. The fact that the repressor is apparently not a protein then raises a serious difficulty since the capacity to form stereospecific complexes with small molecules appears to be a privilege of proteins. If a protein, perhaps an enzyme, is responsible for the specificity, the structure of this protein is presumably determined by a structual gene and mutation in this gene would result in loss of the capacity to be induced (or repressed). Such mutants, which would have precisely predictable properties (they would be pleiotropic, recessive, and they would be complemented by mutants of the other structural genes) have not been encountered in the lactose system, while the possibility that the controlled enzymes themselves (galactosidase or acetylase) play the role of "induction enzyme" is excluded.

It is conceivable that, in the repressible systems which synthesize amino acids, this role is played by enzymes simultaneously responsible for essential functions (e.g. the activating enzymes) whose loss would be lethal, but this seems hardly conceivable in the case of most inducible systems. One possibility which is not excluded by these observations is that the repressor itself synthesizes the "induction protein" and remains thereafter associated with it. Genetic inactivation of the induction enzyme would then be associated with structural alterations of the repressor itself and would generally be expressed as constitutive mutations of the regulator gene.† This possibility is mentioned here only as an illustration of the dilemma which we have briefly analysed, and whose solution will depend upon the chemical identification of the repressor.

(e) *Regulator genes and immunity in temperate phage systems*

One of the most conspicuous examples of the fact that certain genes may be either allowed to express their potentialities, or specifically prohibited from doing so, is the phenomenon of immunity in temperate phage systems (*cf.* Lwoff, 1953; Jacob, 1954; Jacob & Wollman, 1957; Bertani, 1958; Jacob, 1960).

The genetic material of the so-called temperate phages can exist in one of two states within the host cell:

(1) In the *vegetative state*, the phage genome multiplies autonomously. This process, during which all the phage components are synthesized, culminates in the production of infectious phage particles which are released by lysis of the host cell.

(2) In the *prophage state*, the genetic material of the phage is attached to a specific site of the bacterial chromosome in such a way that both genetic elements replicate as a single unit. The host cell is said to be "lysogenic." As long as the phage genome remains in the prophage state, phage particles are not produced. For lysogenic bacteria to produce phage, the genetic material of the phage must undergo a transition from the prophage to the vegetative state. During normal growth of lysogenic bacteria, this event is exceedingly rare. With certain types of prophages, however, the transition can be induced in the whole population by exposure of the culture to u.v. light,

† Such a model could account for the properties of the i^s (dominant) mutant of the regulator gene in the lactose system, by the assumption that in this mutant the repressor remains active, while having lost the capacity to form its associated induction protein.

X-rays or various compounds known to alter DNA metabolism (Lwoff, Siminovitch & Kjeldgaard, 1950; Lwoff, 1953; Jacob, 1954).

The study of "defective" phage genomes, in which a mutation has altered one of the steps required for the production of phage particles, indicates the existence of at least two distinct groups of viral functions, both of which are related to the capacity of synthesizing specific proteins (Jacob, Fuerst & Wollman, 1957). Some "early" functions appear as a pre-requisite for the vegetative multiplication of the phage genome and, at least in virulent phages of the T-even series, it is now known that they correspond to the synthesis of a series of new enzymes (Flaks & Cohen, 1959; Kornberg, Zimmerman, Kornberg & Josse, 1959). A group of "late" functions correspond to the synthesis of the structural proteins which constitute the phage coat. The expression of these different viral functions appears to be in some way co-ordinated by a sequential process, since defective mutations affecting some of the early functions may also result in the loss of the capacity to perform several later steps of phage multiplication (Jacob *et al.*, 1957).

In contrast, the viral functions are not expressed in the prophage state and the protein constituents of the phage coat cannot be detected within lysogenic bacteria. In addition, lysogenic bacteria exhibit the remarkable property of being specifically *immune* to the very type of phage particles whose genome is already present in the cell as prophage. When lysogenic cells are infected with homologous phage particles, these particles absorb onto the cells and inject their genetic material, but the cell survives. The injected genetic material does not express its viral functions: it is unable to initiate the synthesis of the protein components of the coat and to multiply vegetatively. It remains inert and is diluted out in the course of bacterial multiplication (Bertani, 1953; Jacob, 1954).

The inhibition of phage-gene functions in lysogenic bacteria therefore applies not only to the prophage, but also to additional homologous phage genomes. It depends only upon the presence of the prophage (and not upon a permanent alteration, provoked by the prophage, of bacterial genes) since loss of the prophage is both necessary and sufficient to make the bacteria sensitive again.

Two kinds of interpretation may be considered to account for these "immunity" relationships:

(a) The prophage occupies and blocks a *chromosomal* site of the host, specifically required in some way for the vegetative multiplication of the homologous phage.

(b) The prophage produces a *cytoplasmic* inhibitor preventing the completion of some reactions (presumably the synthesis of a particular protein) necessary for the initiation of vegetative multiplication.

A decision between these alternative hypotheses may be reached through the study of persistent diploids, heterozygous for the character lysogeny. A sex factor has been isolated which has incorporated a segment of the bacterial chromosome carrying the genes which control galactose fermentation, *Gal*, and the site of attachment of prophage, λ. Diploid heterozygotes with the structure $Gal^- \lambda^-/F\ Gal^+ \lambda^+$ or $Gal^- \lambda^+/F\ Gal^+ \lambda^-$ are immune against superinfection with phage λ, a result which shows that "immunity" is dominant over "non-immunity" and has a cytoplasmic expression (Jacob, Schaeffer & Wollman, 1960).

The study of transient zygotes formed during conjugation between lysogenic (λ^+) and non-lysogenic (λ^-) cells leads to the same conclusion. In crosses $\male\lambda^+ \times \female\lambda^-$, the transfer of the prophage carried by the male chromosome into the non-immune

recipient results in transition to the vegetative state: multiplication of the phage occurs in the zygotes, which are lysed and release phage particles. This phenomenon is known as "zygotic induction" (Jacob & Wollman, 1956). In the *reverse* cross $\male\lambda^- \times \female\lambda^+$, however, *no zygotic induction occurs*. The transfer of the "non-lysogenic" character carried by the male chromosome into the immune recipient does not bring about the development of the prophage and the zygotes are immune against superinfection with phage λ.

The opposite results obtained in reciprocal crosses of lysogenic by non-lysogenic male and female cells are entirely analogous to the observations made with the lactose system in reciprocal crosses of inducible by non-inducible cells. In both cases, it is evident that the decisive factor is the origin of the *cytoplasm* of the zygote, and the conclusion is inescapable, that the immunity of lysogenic bacteria is due to a cytoplasmic constituent, in the presence of which the viral genes cannot become expressed (Jacob, 1960).

The same two hypotheses which we have already considered for the interpretation of the product of the regulator gene in the lactose system, apply to the cytoplasmic inhibitor insuring immunity in lysogenic bacteria.

(a) The inhibitor is a specific repressor which prevents the synthesis of some early protein(s) required for the initiation of vegetative multiplication.

(b) The inhibitor is an enzyme which destroys a metabolite, normally synthesized by the non-lysogenic cell and specifically required for the vegetative multiplication of the phage.

Several lines of evidence argue against the second hypothesis (Jacob & Campbell, 1959; Jacob, 1960). First, for a given strain of bacteria, many temperate phages are known, each of which exhibits a different immunity pattern. According to the second hypothesis, each of these phages would specifically require for vegetative multiplication a different metabolite normally produced by the non-lysogenic cells, an assumption which appears extremely unlikely. The second argument stems from the fact that, like the repressor of the lactose system, the inhibitor responsible for immunity is synthesized in the presence of chloramphenicol, i.e. in the absence of protein synthesis: when crosses $\male\lambda^+ \times \female\lambda^-$ are performed in the presence of chloramphenicol, no zygotic induction occurs and the prophage is found to segregate normally among recombinants.

In order to explain immunity in lysogenic bacteria, we are led therefore to the same type of interpretation as in the case of adaptive enzyme systems. According to this interpretation, the prophage controls a cytoplasmic repressor, which inhibits specifically the synthesis of one (or several) protein(s) necessary for the initiation of vegetative multiplication. In this model, the introduction of the genetic material of the phage into a non-lysogenic cell, whether by infection or by conjugation, results in a "race" between the synthesis of the specific repressor and that of the early proteins required for vegetative multiplication. The fate of the host-cell, survival with lysogenization or lysis as a result of phage multiplication, depends upon whether the synthesis of the repressor or that of the protein is favoured. Changes in the cultural conditions favoring the synthesis of the repressor such as infection at low temperature, or in the presence of chloramphenicol, would favor lysogenization and *vice versa*. The phenomenon of induction by u.v. light could then be understood, for instance, in the following way: exposure of inducible lysogenic bacteria to u.v. light or X-rays would transiently disturb the regulation system, for example by preventing further synthesis of the repressor. If the repressor is unstable, its concentration inside the cell would

decrease and reach a level low enough to allow the synthesis of the early proteins. Thus the vegetative multiplication would be irreversibly initiated.

The similarity between lysogenic systems and adaptive systems is further strengthened by the genetic analysis of immunity. Schematically, the genome of phage λ appears to involve two parts (see Fig. 5): a small central segment, the *C* region, contains a few determinants which control various functions involved in lysogenization (Kaiser, 1957); the rest of the linkage group contains determinants which govern the "viral functions," i.e. presumably the structural genes corresponding to the different phage proteins. Certain strains of temperate phages which exhibit different immunity patterns are able nevertheless to undergo genetic recombination. The specific immunity pattern segregates in such crosses, proving to be controlled by a small segment "*im*" of the *C* region (Kaiser & Jacob, 1957). In other words, a prophage contains in its *C* region a small segment "*im*" which controls the synthesis of a specific repressor, active on the phage genome carrying a homologous "*im*" segment.

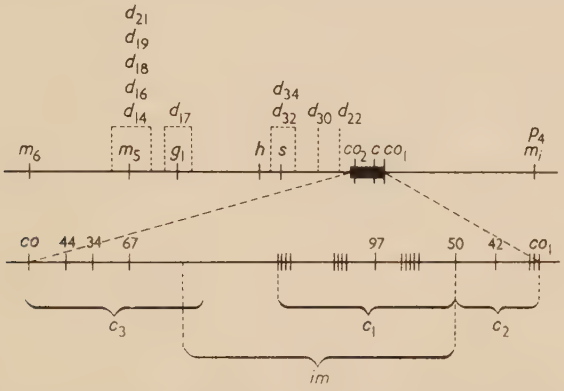


FIG. 5. Diagrammatic representation of the linkage group of the temperate bacteriophage λ. The upper diagram represents the linear arrangement of markers. Symbols refer to various plaque size, plaque type and host-range markers. Symbols *d* refer to various defective mutations. The *C* region represented by a thicker line is enlarged in the lower diagram. The figures correspond to various *C* mutations. The *C* region can be subdivided into three functional units, C_1, C_2 and C_3; the segment controlling immunity is designated *im*.

In the "*im*" region, two types of mutations arise, whose properties are extremely similar to those of the different mutations affecting the regulator genes of adaptive enzyme systems.

(1) Some mutations ($C_I^+ \rightarrow C_I$) result in the complete loss of the capacity for lysogenization in single infection. All the C_I mutations are located in a cluster, in a small part of the "*im*" segment, and they behave as belonging to a single cistron in complementation tests.

In mixed infections with both C_I and C_I^+ phages, double lysogenic clones carrying both C_I and C_I^+ prophages can be recovered. In such clones, single lysogenic cells segregate, which carry the C^+ type alone but never the C_I type alone. These findings indicate that the wild allele is dominant over the mutant C_I alleles and is cytoplasmically expressed, repressing the mutant genome into the prophage state. The properties of the C_I mutations are therefore similar to those of the recessive constitutive mutations of adaptive systems. The evidence suggests that the C_I locus controls the synthesis of the repressor responsible for immunity, and that the C_I mutations correspond to inactivation of this locus, or of its product.

(2) A mutation ($ind^+ \rightarrow ind^-$) has been found which results in the loss of the inducible property of the prophage, i.e. of its capacity to multiply vegetatively upon exposure of lysogenic bacteria to u.v. light, X-rays or chemical inducers. This mutation is located in the C_I segment. The mutant allele *ind*$^-$ is dominant over the wild allele *ind*$^+$ since double lysogenic $\lambda ind^+/\lambda ind^-$ or diploid heterozygotes of structure *Gal*$^-$ $\lambda ind^+/F$ *Gal*$^+\lambda ind^-$ or *Gal* $^-\lambda ind^-/F$ *Gal*$^+\lambda ind^+$ are all non-inducible. In addition, the mutant λind^- exhibits a unique property. If lysogenic bacteria K12 (λ^+) carrying a wild type prophage are exposed to u.v. light, the whole population lyses and releases phage. Infection of such cells with λind^- mutants, either before or immediately after irradiation, completely inhibits phage production and lysis.

The properties of the *ind*$^-$ mutant appear in every respect similar to those of the previously described mutant i^s of the lactose system. The unique properties of the *ind*$^-$ mutants can be explained only by the same type of hypothesis, namely that the mutation *ind*$^-$ affects, quantitatively or qualitatively, the synthesis of the repressor in such a way that more repressor or a more efficient repressor is produced. If this assumption as well as the hypothesis that the C_I mutation results in the loss of the capacity to produce an active repressor, are correct, the double mutants $C_I ind^-$ should have lost the capacity of inhibiting phage multiplication upon infection of wild type lysogenic cells. This is actually what is observed. It is evident that the properties of the *ind*$^-$ mutant cannot be accounted for by the assumption that the C_I locus controls the synthesis of a metabolite-destroying enzyme (Jacob & Campbell, 1959).

In summary, the analysis of lysogenic systems reveals that the expression of the viral genes in these systems is controlled by a cytoplasmic repressor substance, whose synthesis is governed by one particular "regulator" gene, belonging to the viral genome. The identity of the proteins whose synthesis is thus repressed is not established, but it seems highly probable that they are "early" enzymes which initiate the whole process of vegetative multiplication. With the (important) limitation that they are sensitive to entirely different types of inducing conditions, the phage repression systems appear entirely comparable to the systems involved in enzymatic adaptation.

4. The Operator and the Operon

(a) *The operator as site of action of the repressor*

In the preceding section we have discussed the evidence which shows that the transfer of information from structural genes to protein is controlled by specific repressors synthesized by specialized regulator genes. We must now consider the next problem, which is the site and mode of action of the repressor.

In regard to this problem, the most important property of the repressor is its characteristic pleiotropic specificity of action. In the lactose system of *E. coli*, the repressor is both *highly specific* since mutations of the *i* gene do not affect any other system, and *pleiotropic* since both galactosidase and acetylase are affected simultaneously and quantitatively to the same extent, by such mutations.

The specificity of operation of the repressor implies that it acts by forming a stereospecific combination with a constituent of the system possessing the proper (complementary) molecular configuration. Furthermore, it must be assumed that the flow of information from gene to protein is interrupted when this element is combined with

the repressor. This controlling element we shall call the "*operator*" (Jacob & Monod, 1959). We should perhaps call attention to the fact that, once the existence of a specific repressor is considered as established, the existence of an operator element defined as above follows necessarily. Our problem, therefore, is not whether an operator exists, but where (and how) it intervenes in the system of information transfer.

An important prediction follows immediately from the preceding considerations. Under any hypothesis concerning the nature of the operator, its specific complementary configuration must be genetically determined; therefore it could be affected by mutations which would alter or abolish its specific affinity for the repressor, without necessarily impairing its activity as initiator of information-transfer. Such mutations would result in *constitutive* synthesis of the protein or proteins. These mutations would define an "operator locus" which should be genetically distinct from the regulator gene (i.e. its mutations should not behave as alleles of the regulator); the most distinctive predictable property of such mutants would be that the constitutive allele should be *dominant* over the wild type since, again under virtually any hypothesis, the presence in a diploid cell of repressor-sensitive operators would not prevent the operation of repressor-insensitive operators.

(b) *Constitutive operator mutations*

Constitutive mutants possessing the properties predicted above have so far been found in two repressor-controlled systems, namely the phage λ and *Lac* system of *E. coli*.

In the case of phage λ, these mutants are characterized, and can be easily selected, by the fact that they develop vegetatively in immune bacteria, lysogenic for the wild type. This characteristic property means that these mutants (v) are *insensitive* to the repressor present in lysogenic cells. When, in fact, lysogenic cells are infected with these mutant particles, the development of the wild type prophage is induced, and the resulting phage population is a mixture of v and v^+ particles. This is expected, since presumably the initiation of prophage development depends only on the formation of one or a few "early" enzyme-proteins, which are supplied by the virulent particle (Jacob & Wollman, 1953).

In the *Lac* system, dominant constitutive (o^c) mutants have been isolated by selecting for constitutivity in cells diploid for the *Lac* region, thus virtually eliminating the recessive (i^-) constitutive mutants (Jacob *et al.*, 1960*a*). By recombination, the o^c mutations can be mapped in the *Lac* region, between the i and the z loci, the order being (*Pro*) ***yzoi*** (*Ad*) (see Fig. 3). Some of the properties of these mutants are summarized in Table 3. To begin with, let us consider only the effects of this mutation on galactosidase synthesis. It will be noted that in the absence of inducer, these organisms synthesize 10 to 20% of the amount of galactosidase synthesized by i^- mutants, i.e. about 100 to 200 times more than uninduced wild type cells (Table 3, lines 3 and 7). In the presence of inducer, they synthesize maximal amounts of enzyme. They are therefore only partially constitutive (except however under conditions of starvation, when they form maximum amounts of galactosidase in the absence of inducer (Brown, unpublished results)). The essential point however is that the enzyme is synthesized constitutively by diploid cells of constitution o^c/o^+ (see Table 3). The o^c allele therefore is "dominant."

If the constitutivity of the o^c mutant results from a loss of sensitivity of the operator to the repressor, the o^c organisms should also be insensitive to the presence of the

altered repressor synthesized by the i^s (dominant) allele of the i^+ gene (see page 331). That this is indeed the case, as shown by the constitutive behavior of diploids with the constitution $i^s o^+ / F i^+ o^c$ (see Table 3, line 12), is a very strong confirmation of the interpretation of the effects of *both* mutations (i^s and o^c). In addition, and as one would expect according to this interpretation, o^c mutants frequently arise as lactose positive "revertants" in populations of i^s cells (see p. 332).

TABLE 3

Synthesis of galactosidase, cross-reacting material (CRM), and galactoside-transacetylase by haploid and heterozygous diploid operator mutants

Strain No.	Genotype	Galactosidase		Cross-reacting material	
		Non-induced	Induced	Non-induced	Induced
1	o^+z^+	<0·1	100	—	—
2	$o^+z^+/Fo^+z_1^-$	<0·1	105	<1	310
3	o^cz^+	15	90	—	—
4	o^+z^+/Fo^cz_1	<0·1	90	30	180
5	$o^+z_1^-/Fo^cz^+$	90	250	<1	85

		Galactosidase		Galactoside-transacetylase	
		Non-Induced	Induced	Non-induced	Induced
6	$o^+z^+y^+$	<0·1	100	<1	100
7	$o^cz^+y^+$	25	95	15	110
8	$o^+z^+y_U^-/Fo^cz^+y^+$	70	220	50	160
9	$o^+z_1^-y^+/Fo^cz^+y_U^-$	180	440	<1	220
10	$i^+o_{84}^\circ z^+y^+$	<0·1	<0·1	<1	<1
11	$i^+o_{84}^\circ z^+y^+/Fi^-o^+z^+y^+$	1	260	2	240
12	$i^so^+z^+y^+/Fi^+o^cz^+y^+$	190	210	150	200

Bacteria are grown in glycerol as carbon source and induced when stated, with *iso*propyl-thiogalactoside, 10^{-4} M. Values of galactosidase and acetylase are given as a percentage of those observed with induced wild type. Values of CRM are expressed as antigenic equivalents of galactosidase. Note that the proteins corresponding to the alleles carried by the sex factor are often produced in greater amount than that observed with induced haploid wild type. This is presumably due to the existence of several copies of the *F-Lac* factor per chromosome. In o^c mutants, haploid or diploid, the absolute values of enzymes produced, especially in the non-induced cultures varies greatly from day to day depending on the conditions of the cultures.

We therefore conclude that the $o^+ \rightarrow o^c$ mutations correspond to a modification of the specific, repressor-accepting, structure of the operator. This identifies the operator locus, i.e. the genetic segment responsible for the structure of the operator, but not the operator itself.

(c) *The operon*

Turning now to this problem, we note that the o^c mutation (like the i^- mutation) is pleiotropic: it affects simultaneously and quantitatively to the same extent, the synthesis of galactosidase and acetylase (see Table 3, lines 7 and 8). The structure of the operator, or operators, which controls the synthesis of the two proteins, therefore, is controlled by a single determinant.†

Two alternative interpretations of this situation must be considered:

(a) A single operator controls an *integral* property of the *z-y* genetic segment, or of its cytoplasmic product.

(b) The specific product of the operator locus is able to associate in the cytoplasm, with the products of the *z* and *y* cistrons, and thereby governs the expression of both structural genes.

The second interpretation implies that mutations of the operator locus should behave as belonging to a cistron *independent* of both the *z* and *y* cistrons. The first interpretation requires, on the contrary, that these mutations behave functionally as if they *belonged to both cistrons simultaneously*. These alternative interpretations can therefore be distinguished without reference to any particular physical model of operator action by testing for the *trans* effect of *o* alleles, that is to say for the constitutive *vs* inducible expression of the two structural genes in o^+/o^c diploids, heterozygous for one or both of these structural genes.

The results obtained with diploids of various structures are shown in Table 3. We may first note that in diploids of constitution $o^+z^+/Fo^cz_1^-$ or $o^+z_1^-/Fo^cz^+$ (lines 4 and 5), both the normal galactosidase produced by the z^+ allele and the altered protein (CRM) produced by the z_1^- allele are formed in the presence of inducer, while in the *absence* of inducer, *only the protein corresponding to the* z *allele in position cis to the* o^c *is produced.* The o^c therefore has no effect on the *z* allele in position *trans*. Or putting it otherwise: the expression of the *z* allele attached to an o^+ remains fully repressor-sensitive even in the presence of an o^c in position *trans*. The *o* locus might be said to behave as belonging to the same cistron as the *z* markers. But as we know already, the o^c mutation is equally effective towards the acetylase which belongs to a cistron independent of *z*, and not adjacent to the operator locus. The results shown in Table 3, lines 8 and 9, confirm that the $o \rightarrow y$ relationship is the same as the $o \rightarrow z$ relationship, that is, the effect of the o^c allele extends *exclusively* to the *y* allele in the *cis* position. For instance, in the diploid $o^+z^-y^+/Fo^cz^+y_U^-$ the galactosidase is constitutive and the acetylase is inducible, while in the diploid $o^+z^+y_U^-/Fo^cz^+y^+$ both enzymes are constitutive.

These observations, predicted by the first interpretation, are incompatible with the second and lead to the conclusion that the operator governs an integral property of the genetic segment *ozy*, or of its cytoplasmic product (Jacob *et al.*, 1960*a*; Képès, Monod & Jacob, 1961).

This leads to another prediction. Certain mutations of the *o* segment could modify the operator in such a way as to inactivate the whole *ozy* segment resulting in the loss of the capacity to synthesize *both* galactosidase and permease.

These "o^o" mutants would be *recessive* to o^+ or o^c, and they would *not* be complemented either by $o^+z^+y^-$ or by $o^+z^-y^+$ mutants. Several point-mutants, possessing

† Let us recall again that no *non-pleiotropic* constitutive mutants of any type have been isolated in this system, in spite of systematic screening for such mutants.

precisely these properties, have been isolated (Jacob *et al.*, 1960*a*). They all map very closely to o^c, as expected (see Fig. 3). It is interesting to note that in these mutants the i^+ gene is functional (Table 3, line 11), which shows clearly, not only that the *i* and *o* mutants are not alleles, but that the *o* segment, while governing the expression of the *z* and *y* genes, does not affect the expression of the regulator gene.

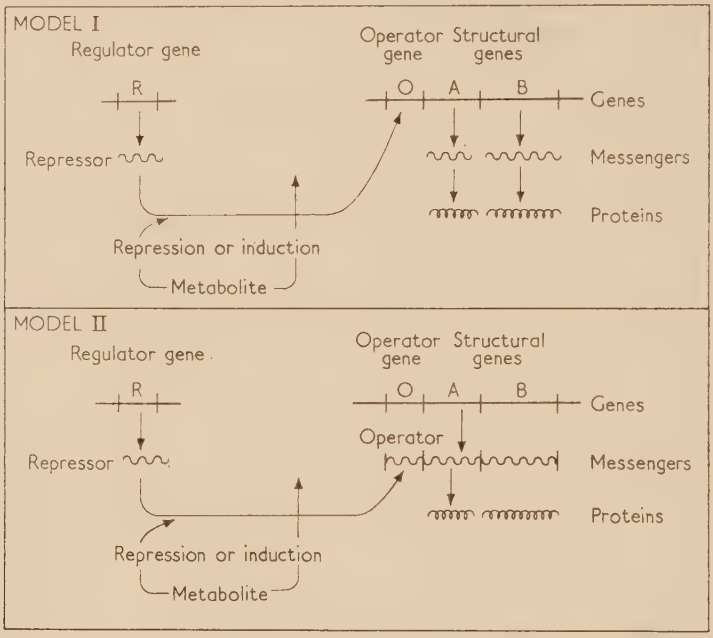


Fig. 6. Models of the regulation of protein synthesis.

In conclusion, the integral or *co-ordinate* expression of the *ozy* genetic segment signifies that the operator, which controls this expression, is and remains attached (see Fig. 6):

(a) either to the genes themselves (Fig. 6, I),

(b) or to the cytoplasmic messenger of the linked *z* and *y* genes which must then be assumed to form a single, integral, particle corresponding to the structure of the whole *ozy* segment, and functioning as a whole (Fig. 6, II).

In the former case, *the operator would in fact be identical with the* o *locus* and it would govern directly the activity of the genes, i.e. the synthesis of the structural messengers.

Both of these models are compatible with the observations which we have discussed so far. We shall return in the next section to the question whether the operator, i.e. the site of specific interaction with the repressor, is genetic or cytoplasmic. In either case, the *ozy* segment, although containing at least two independent structural genes, governing two independent proteins, behaves as a *unit* in the transfer of information. This *genetic unit of co-ordinate expression* we shall call the "*operon*" (Jacob *et al.*, 1960*a*).

The existence of such a unit of genetic expression is proved so far only in the case of the *Lac* segment. As we have already seen, the *v* mutants of phage λ, while illustrating the existence of an operator in this system, do not define an operon (because the

number and the functions of the structural genes controlled by this operator are unknown). However, many observations hitherto unexplained by or even conflicting with classical genetic theory, are immediately accounted for by the operon theory. It is well known that, in bacteria, the genes governing the synthesis of different enzymes sequentially involved in a metabolic pathway are often found to be extremely closely linked, forming a cluster (Demerec, 1956). Various not very convincing speculations have been advanced to account for this obvious correlation of genetic structure and biochemical function (see Pontecorvo, 1958). Since it is now established that simultaneous induction or repression also generally prevails in such metabolic sequences, it seems very likely that the gene clusters represent units of co-ordinate expression, i.e. operons.

We have already mentioned the fact that two inducible enzymes sequentially involved in the metabolism of galactose by *E. coli*, galactokinase and UDP-galactose-transferase, are simultaneously induced by galactose, or by the gratuitous inducer D-fucose (Buttin, 1961). The genes which control specifically the synthesis of these enzymes, i.e. presumably the structural genes, are closely linked, forming a cluster on the *E. coli* chromosome. (Kalckar, Kurahashi & Jordan, 1959; Lederberg, 1960; Yarmolinsky & Wiesmeyer, 1960; Adler, unpublished results.) Certain point-mutations which occur in this chromosome segment abolish the capacity to synthesize both enzymes. These pleiotropic loss mutations are not complemented by any one of the specific (structural) loss mutations, an observation which is in apparent direct conflict with the one-gene one-enzyme hypothesis. These relationships are explained and the conflict is resolved if it is assumed that the linked structural genes constitute an operon controlled by a single operator and that the pleiotropic mutations are mutations of the operator locus.

We have also already discussed the system of simultaneous repression which controls the synthesis of the enzymes involved in histidine synthesis in *Salmonella*. This system involves eight or nine reaction steps. The enzymes which catalyse five of these reactions have been identified. The genes which individually determine these enzymes form a closely linked cluster on the *Salmonella* chromosome. Mutations in each of these genes result in a loss of capacity to synthesize a single enzyme; however, certain mutations at one end of the cluster abolish the capacity to synthesize all the enzymes simultaneously, and these mutations are not complemented by any one of the specific mutations (Ames, Garry & Herzenberg, 1960; Hartman, Loper & Serman, 1960). It will be recalled that the relative rates of synthesis of different enzymes in this sequence are constant under any set of conditions (see p. 327). All these remarkable findings are explained if it is assumed that this cluster of genes constitutes an operon, controlled by an operator associated with the *g* cistron.

The rule that genes controlling metabolically sequential enzymes constitute genetic clusters does not apply, in general, to organisms other than bacteria (Pontecorvo, 1958). Nor does it apply to all bacterial systems, even where simultaneous repression is known to occur and to be controlled by a single regulator gene, as is apparently the case for the enzymes of arginine biosynthesis. In such cases, it must be supposed that several identical or similar operator loci are responsible for sensitivity to repressor of each of the independent information-transfer systems.

It is clear that when an operator controls the expression of only a single structural cistron, the concept of the operon does not apply, and in fact there are no conceivable genetic-biochemical tests which could identify the operator-controlling genetic

segment as distinct from the structural cistron itself.† One may therefore wonder whether it will be possible experimentally to extend this concept to dispersed (as opposed to clustered) genetic systems. It should be remarked at this point that many enzyme proteins are apparently made up of two (or more) different polypeptide chains. It is tempting to predict that such proteins will often be found to be controlled by two (or more) adjacent and co-ordinated structural cistrons, forming an operon.

5. The Kinetics of Expression of Structural Genes, and the Nature of the Structural Message

The problem we want to discuss in this section is whether the repressor-operator system functions at the genetic level by governing the *synthesis* of the structural message or at the cytoplasmic level, by controlling the protein-synthesizing *activity* of the messenger (see Fig. 6). These two conceivable models we shall designate respectively as the "genetic operator model" and the "cytoplasmic operator model."

The existence of units of co-ordinate expression involving several structural genes appears in fact difficult to reconcile with the cytoplasmic operator model, if only because of the size that the cytoplasmic unit would have to attain. If we assume that the message is a polyribonucleotide and take a coding ratio of 3, the "unit message" corresponding to an operon governing the synthesis of three proteins of average (monomeric) molecular weight 60,000 would have a molecular weight about $1{\cdot}8 \times 10^6$; we have seen that operons including up to 8 structural cistrons may in fact exist. On the other hand, RNA fractions of *E. coli* and other cells do not appear to include polyribonucleotide molecules of molecular weight exceeding 10^6.

This difficulty is probably not insuperable; and this type of argument, given the present state of our knowledge, cannot be considered to eliminate the cytoplasmic operator model, even less to establish the validity of the genetic model. However, it seems more profitable tentatively to adopt the genetic model and to see whether some of the more specific predictions which it implies are experimentally verified.

The most immediate and also perhaps the most striking of these implications is that the structural message must be carried by a very short-lived intermediate both rapidly formed and rapidly destroyed during the process of information transfer. This is required by the kinetics of induction. As we have seen, the addition of inducer, or the removal of co-repressor, provokes the synthesis of enzyme at maximum rate within a matter of a few minutes, while the removal of inducer, or the addition of co-repressor interrupts the synthesis within an equally short time. Such kinetics are incompatible with the assumption that the repressor-operator interaction controls the rate of synthesis of *stable* enzyme-forming templates (Monod, 1956, 1958). Therefore, if the genetic operator model is valid, one should expect the kinetics of structural gene expression to be *essentially the same* as the kinetics of induction: injection of a "new" gene into an otherwise competent cell should result in virtually immediate synthesis of the corresponding protein at maximum rate; while removal of the gene should be attended by concomitant cessation of synthesis.

† It should be pointed out that the operational distinction between the operator locus and the structural cistron to which it is directly adjacent rests exclusively on the fact that the operator mutations affect the synthesis of several proteins governed by linked cistrons. This does not exclude the possibility that the operator locus is actually *part* of the structural cistron to which it is "adjacent." If it were so, one might expect certain constitutive operator mutations to involve an alteration of the structure of the protein governed by the "adjacent" cistron. The evidence available at present is insufficient to confirm or eliminate this assumption.

(a) *Kinetics of expression of the galactosidase structural gene*

Additions and removals of genes to and from cells are somewhat more difficult to perform than additions or removals of inducer. However, it can be done. Gene injection without cytoplasmic mixing occurs in the conjugation of *Hfr* male and F^- female *E. coli*. In a mixed male and female population the individual pairs do not all mate at the same time, but the distribution of times of injection of a *given* gene can be rather accurately determined by proper genetic methods. The injection of the z^+ (galactosidase) gene from male cells into galactosidase-negative (z^-) female cells is rapidly followed by enzyme synthesis within zygotes (cf. p. 332). When the rate of enzyme synthesis in the population is expressed as a function of time, taking into

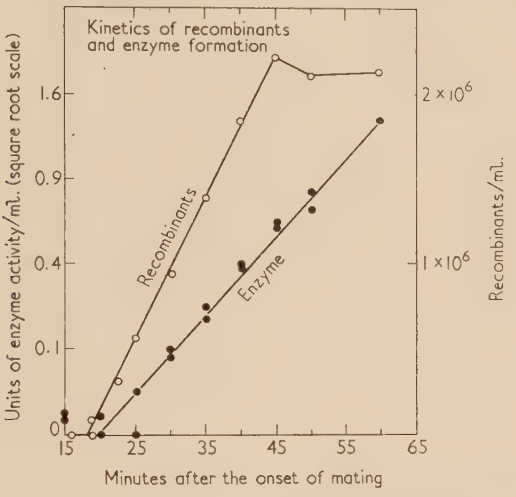


FIG. 7. Kinetics of enzyme production by merozygotes formed by conjugation between inducible galactosidase-positive males and constitutive galactosidase-negative females. Conditions are such that only the zygotes can form enzyme. Increase in the number of z^+ containing zygotes is determined by counting recombinants on adequate selective medium. Formation of enzyme is followed by enzyme activity measurements on the total population. It is seen that the enzyme increases linearly with the square of time. Since the zygote population increases linearly with time, it is apparent that the rate of enzyme synthesis per zygote is constant from the time of penetration of the z^+ gene. (From Riley *et al.*, 1960.)

account the increase with time of the number of z^+ containing zygotes, it is found (see Fig. 7):

(1) that enzyme synthesis begins within two minutes of the penetration of the z^+ gene;

(2) that the rate per zygote is constant and maximum over at least the first 40 min following penetration (Riley, Pardee, Jacob & Monod, 1960).

These observations indicate that the structural messenger is very rapidly formed by the z^+ gene, and does not accumulate. This could be interpretated in one of two ways:

(a) the structural messenger is a short-lived intermediate;

(b) the structural messenger is stable, but the gene rapidly forms a limited number of messenger molecules, and thereafter stops functioning.

If the second assumption is correct, removal of the gene after the inception of enzyme synthesis should not prevent the synthesis from continuing. This possibility is tested by the "removal" experiment, which is performed by loading the male

chromosome with ^{32}P before injection. Following injection (into unlabelled female cells), ample time (25 min) is allowed for expression of the z^+ gene, before the zygotes are frozen to allow ^{32}P decay for various lengths of time. The rate of galactosidase synthesis by the population is determined immediately after thawing. It is found to decrease sharply as a function of the fraction of ^{32}P atoms decayed. If a longer period of time (110 min) is allowed for expression before freezing, no decrease in either enzyme-forming capacity or in viability of the z^+ marker are observed. This is to be expected, since by that time most of the z^+ genes would have replicated, and this observation provides an internal control showing that no indirect effects of ^{32}P disintegrations are involved.

This experiment therefore indicates that even after the z^+ gene has become expressed its integrity is required for enzyme synthesis to continue, as expected if the messenger molecule is a short-lived intermediate (Riley *et al.*, 1960).

The interpretation of both the injection and the removal experiment rests on the assumption that the observed effects are not due to (stable) cytoplasmic messenger molecules introduced with the genetic material, during conjugation. As we have already noted, there is strong evidence that no cytoplasmic transfer, even of small molecules, occurs during conjugation. Furthermore, if the assumption were made that enzyme synthesis in the zygotes is due to pre-formed messenger molecules rather than to the activity of the gene, it would be exceedingly difficult to account for both (a) the very precise coincidence in time between inception of enzyme synthesis and entry of the gene (in the injection experiment) and (b) the parallel behaviour of enzyme-forming capacity and genetic viability of the z^+ gene (in the removal experiment).

These experiments therefore appear to show that the kinetics of expression of a structural gene are entirely similar to the kinetics of induction-repression, as expected if the operator controls the activity of the gene in the synthesis of a short-lived messenger, rather than the activity of a ready-made (stable) messenger molecule in synthesizing protein.

It is interesting at this point to recall the fact that infection of *E. coli* with virulent (ϕII, T2, T4) phage is attended within 2 to 4 minutes by inhibition of *bacterial* protein synthesis, including in particular β-galactosidase (Cohen, 1949; Monod & Wollman, 1947; Benzer, 1953). It is known on the other hand that phage-infection results in rapid visible lysis of bacterial nuclei, while no major destruction of pre-formed bacterial RNA appears to occur (Luria & Human, 1950). It seems very probable that the inhibition of specific bacterial protein synthesis by virulent phage is due essentially to the depolymerization of bacterial DNA, and this conclusion also implies that the integrity of bacterial genes is required for continued synthesis of bacterial protein. In confirmation of this interpretation, it may be noted that infection of *E. coli* by phage λ, which does not result in destruction of bacterial nuclei, allows β-galactosidase synthesis to continue almost to the time of lysis (Siminovitch & Jacob, 1952).

(b) *Structural effects of base analogs*

An entirely different type of experiment also leads to the conclusion that the structural messenger is a short-lived intermediate and suggests, furthermore, that this intermediate is a ribonucleotide. It is known that certain purine and pyrimidine analogs are incorporated by bacterial cells into ribo- and deoxyribonucleotides, and it has been found that the synthesis of protein, or of some proteins, may be inhibited in the presence of certain of these analogs. One of the mechanisms by which these effects

could be explained may be that certain analogs are incorporated into the structural messenger. If so, one might hope to observe that the molecular structure of specific proteins formed in the presence of an analog is modified. It has in fact been found that the molecular properties of β-galactosidase and of alkaline phosphatase synthesized by *E. coli* in the presence of 5-fluorouracil (5FU) are strikingly altered. In the case of β-galactosidase, the ratio of enzyme activity to antigenic valency is decreased by 80%. In the case of alkaline phosphatase, the rate of thermal inactivation (of this normally highly heat-resistant protein) is greatly increased (Naono & Gros, 1960*a*,*b*; Bussard, Naono, Gros & Monod, 1960).

It can safely be assumed that such an effect cannot result from the mere presence of 5FU in the cells, and must reflect incorporation of the analog into a constituent involved in some way in the information transfer system. Whatever the identity of this constituent may be, the kinetics of the effect must in turn reflect the kinetics of 5FU incorporation into this constituent. The most remarkable feature of the 5FU effect is that it is almost immediate, in the sense that abnormal enzyme is synthesized almost from the time of addition of the analog, and that the degree of abnormality of the molecular population thereafter synthesized does not increase with time. For instance, in the case of galactosidase abnormal enzyme is synthesized within 5 min of addition of the analog, and the ratio of enzyme activity to antigenic valency remains constant thereafter. In the case of alkaline phosphatase, the thermal inactivation curve of the abnormal protein synthesized in the presence of 5FU is monomolecular, showing the molecular population to be *homogeneously* abnormal rather than made up of a mixture of normal and abnormal molecules. It is clear that if the constituent responsible for this effect were stable, one would expect the population of molecules made in the presence of 5FU to be heterogeneous, and the fraction of abnormal molecules to increase progressively. It follows that the responsible constituent must be formed, and also must decay, very rapidly.

Now it should be noted that, besides the structural gene-synthesized messenger, the information transfer system probably involves other constituents responsible for the correct translation of the message, such as for instance the RNA fractions involved in amino acid transfer. The 5FU effect could be due to incorporation into one of these fractions rather than to incorporation into the messenger itself. However, the convergence of the results of the different experiments discussed above strongly suggests that the 5FU effect does reflect a high rate of turnover of the messenger itself.

(c) *Messenger RNA*

Accepting tentatively these conclusions, let us then consider what properties would be required of a cellular constituent, to allow its identification with the structural messenger. These qualifications based on general assumptions, and on the results discussed above, would be as follows:

(1) The "candidate" should be a polynucleotide.

(2) The fraction would presumably be very heterogeneous with respect to molecular weight. However, assuming a coding ratio of 3, the average molecular weight would not be lower than 5×10^5.

(3) It should have a base composition reflecting the base composition of DNA.

(4) It should, at least temporarily or under certain conditions, be found associated with ribosomes, since there are good reasons to believe that ribosomes are the seat of protein synthesis.

(5) It should have a very high rate of turnover and in particular it should saturate with 5FU within less than about 3 min.

It is immediately evident that none of the more classically recognized cellular RNA fractions meets these very restrictive qualifications. Ribosomal RNA, frequently assumed to represent the "template" in protein synthesis, is remarkably homogeneous in molecular weight. Its base composition is similar in different species, and does not reflect the variations in base ratios found in DNA. Moreover it appears to be entirely stable in growing cells (Davern & Meselson, 1960). It incorporates 5FU only in proportion to net increase.

Transfer RNA, or (sRNA) does not reflect DNA in base composition. Its average molecular weight is much lower than the 5×10^5 required for the messenger. Except perhaps for the terminal adenine and cytidine, its rate of incorporation of bases, including in particular 5FU, is not higher than that of ribosomal RNA.

However, a small fraction of RNA, first observed by Volkin & Astrachan (1957) in phage infected *E. coli*, and recently found to exist also in normal yeasts (Yčas & Vincent, 1960) and coli (Gros, *et al.*, 1961), does seem to meet all the qualifications listed above.

This fraction (which we shall designate "messenger RNA" or M-RNA) amounts to only about 3% of the total RNA; it can be separated from other RNA fractions by column fractionation or sedimentation (Fig. 8). Its average sedimentation velocity coefficient is 13, corresponding to a minimum molecular weight of 3×10^5, but since the molecules are presumably far from spherical, the molecular weight is probably much higher. The rate of incorporation of ^{32}P, uracil or 5FU into this fraction is extremely rapid: half saturation is observed in less than 30 sec, indicating a rate of synthesis several hundred times faster than any other RNA fraction. Its half life is also very short, as shown by the disappearance of radioactivity from this fraction in pre-labelled cells. At high concentrations of Mg^{2+} (0·005 M) the fraction tends to associate with the 70s ribosomal particles, while at lower Mg^{2+} concentrations it sediments independently of the ribosomal particles (Gros *et al.*, 1961).

The striking fact, discovered by Volkin & Astrachan, that the base-composition of this fraction in T2-infected cells reflects the base composition of *phage* (rather than bacterial) DNA, had led to the suggestion that it served as a precursor of phage DNA. The agreement between the properties of this fraction and the properties of a short-lived structural messenger suggests that, in phage infected cells as well as in normal cells, this fraction served in fact in the transfer of genetic information from phage DNA to the protein synthesizing centers. This assumption implies that the same protein-forming centers which, in uninfected cells, synthesize bacterial protein, also serve in infected cells to synthesize phage protein according to the new structural information provided by phage DNA, *via* M-RNA. This interpretation is strongly supported by recent observations made with T4 infected *E. coli*. (Brenner, Jacob & Meselson, 1961).

Uninfected cells of *E. coli* were grown in the presence of ^{15}N. They were then infected and resuspended in ^{14}N medium. Following infection, they were exposed to short pulses of ^{32}P or ^{35}S, and the ribosomes were analysed in density gradients. It was found:

(1) that no detectable amounts of ribosomal RNA were synthesized after infection;

(2) labelled M-RNA formed *after* infection became associated with unlabelled ribosomal particles formed *before* infection;

(3) newly formed (i.e. phage-determined) protein, identified by its ^{35}S content, was found associated with the 70s particles before it appeared in the soluble protein fraction.

These observations strongly suggest that phage protein is synthesized by *bacterial* ribosomes formed before infection and associated with *phage-determined* M-RNA. Since the structural information for phage protein could not reside in the bacterial ribosomes, it must be provided by the M-RNA fraction.

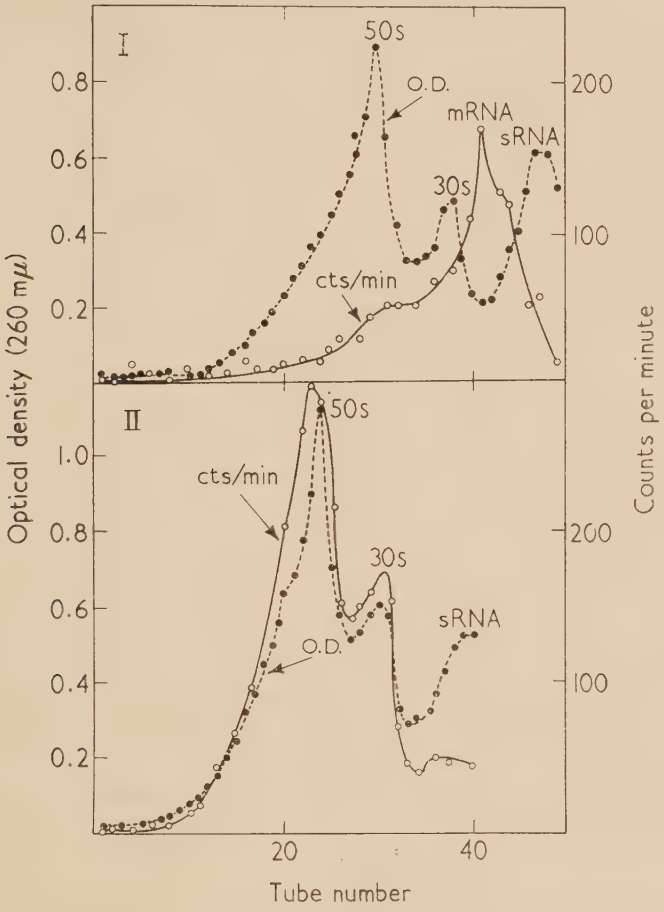


FIG. 8. Incorporation and turnover of uracil in messenger RNA. *E. coli* growing exponentially in broth were incubated for 5 sec with [^{14}C]-uracil. The bacteria were centrifuged, washed and resuspended in the original volume of the same medium containing 100-fold excess of [^{12}C]-uracil. Half the bacteria were then harvested and frozen (I) and the remainder were incubated for 15 min at 37°C (II) prior to harvesting and freezing. The frozen samples were ground with alumina and extracted with tris buffer (2-amino-2 hydroxymethylpropane-1:3-diol) containing 10^{-4}M-Mg, treated with DNase and applied to a sucrose gradient. After 3 hr, sequential samples were taken for determination of radioactivity and absorption at 260 mμ. It may be seen (part I) that after 5 sec, M-RNA is the only labelled fraction, and that subsequently (part II) uracil incorporated into M-RNA is entirely renewed. (From Gros *et al.*, 1961.)

Finally, the recent experiments of Lamfrom (1961) independently repeated by Kruh, Rosa, Dreyfus & Schapira (1961) have shown directly that species specificity in the synthesis of haemoglobin is determined by a "soluble" RNA-containing fraction rather than by the ribosomal fraction. Lamfrom used reconstructed systems, containing ribosomes from one species (rabbit) and soluble fractions from another (sheep) and found that the haemoglobin formed *in vitro* by these systems belonged in part to the

type characteristic of the species used to prepare the *soluble* fraction. It is not, of course, positively proved that *inter-specific* differences in haemoglobin structure are gene-determined rather than cytoplasmic, but the assumption seems safe enough. In any case, Lamfrom's experiment proves beyond doubt that the ribosomes cannot be considered to determine entirely (if at all) the specific structure of proteins.

We had stated the problem to be discussed in this section as the choice between the genetic operator model and the cytoplasmic operator model. The adoption of the genetic operator model implies, as we have seen, some very distinctive and specific predictions concerning the behaviour of the intermediate responsible for the transfer of information from gene to protein. These predictions appear to be borne out by a considerable body of evidence which leads actually to a tentative identification of the intermediate in question with one particular RNA fraction. Even if this identification is confirmed by direct experiments, it will remain to be proved, also by direct experiments, that the synthesis of this "M-RNA" fraction is controlled at the genetic level by the repressor-operator interaction.

6. Conclusion

A convenient method of summarizing the conclusions derived in the preceding sections of this paper will be to organize them into a model designed to embody the main elements which we were led to recognize as playing a specific role in the control of protein synthesis; namely, the structural, regulator and operator genes, the operon, and the cytoplasmic repressor. Such a model could be as follows:

The molecular structure of proteins is determined by specific elements, the *structural genes*. These act by forming a cytoplasmic "transcript" of themselves, the structural messenger, which in turn synthesizes the protein. The synthesis of the messenger by the structural gene is a sequential replicative process, which can be initiated only at certain points on the DNA strand, and the cytoplasmic transcription of several, linked, structural genes may depend upon a single initiating point or *operator*. The genes whose activity is thus co-ordinated form an *operon*.

The operator tends to combine (by virtue of possessing a particular base sequence) specifically and reversibly with a certain (RNA) fraction possessing the proper (complementary) sequence. This combination blocks the initiation of cytoplasmic transcription and therefore the formation of the messenger by the structural genes in the whole operon. The specific "repressor" (RNA?), acting with a given operator, is synthesized by a *regulator gene*.

The repressor in certain systems (inducible enzyme systems) tends to combine specifically with certain specific small molecules. The combined repressor has no affinity for the operator, and the combination therefore results in *activation of the operon*.

In other systems (repressible enzyme systems) the repressor by itself is inactive (i.e. it has no affinity for the operator) and is activated only by combining with certain specific small molecules. The combination therefore leads to *inhibition of the operon*.

The structural messenger is an unstable molecule, which is destroyed in the process of information transfer. The rate of messenger synthesis, therefore, in turn controls the rate of protein synthesis.

This model was meant to summarize and express conveniently the properties of the different factors which play a specific role in the control of protein synthesis. In

order concretely to represent the functions of these different factors, we have had to introduce some purely speculative assumptions. Let us clearly discriminate the experimentally established conclusions from the speculations:

(1) The most firmly grounded of these conclusions is the existence of *regulator* genes, which control the rate of information-transfer from *structural* genes to proteins, without contributing any information to the proteins themselves. Let us briefly recall the evidence on this point: mutations in the structural gene, which are reflected as alterations of the protein, do not alter the regulatory mechanism. Mutations that alter the regulatory mechanism do not alter the protein and do not map in the structural genes. Structural genes obey the one-gene one-protein principle, while regulator genes may affect the synthesis of several different proteins.

(2) That the regulator gene acts *via* a specific cytoplasmic substance whose effect is to *inhibit* the expression of the structural genes, is equally clearly established by the *trans* effect of the gene, by the different properties exhibited by genetically identical zygotes depending upon the origin of their cytoplasm, and by the fact that absence of the regulator gene, or of its product, results in uncontrolled synthesis of the protein at maximum rates.

(3) That the product of the regulator gene acts directly as a *repressor* (rather than indirectly, as antagonist of an endogenous inducer or other activator) is proved in the case of the *Lac* system (and of the λ lysogenic systems) by the properties of the dominant mutants of the regulator.

(4) The chemical identification of the repressor as an RNA fraction is a logical assumption based only on the *negative* evidence which indicates that it is not a protein.

(5) The existence of an operator, defined as the site of action of the repressor, is deduced from the existence and specificity of action of the repressor. The identification of the operator with the genetic segment which controls sensitivity to the repressor, is strongly suggested by the observation that a *single* operator gene may control the expression of *several adjacent structural genes*, that is to say, by the demonstration of the *operon* as a co-ordinated unit of genetic expression.

The assumption that the operator represents an initiating point for the cytoplasmic transcription of several structural genes is a pure speculation, meant only as an illustration of the fact that the operator controls an integral property of the group of linked genes which form an operon. There is at present no evidence on which to base any assumption on the molecular mechanisms of the operator.

(6) The assumptions made regarding the interaction of the repressor with inducers or co-repressors are among the weakest and vaguest in the model. The idea that specific coupling of inducers to the repressor could result in inactivation of the repressor appears reasonable enough, but it raises a difficulty which we have already pointed out. Since this reaction between repressor and inducer must be stereospecific (for both) it should presumably require a specific enzyme; yet no evidence, genetic or biochemical, has been found for such an enzyme.

(7) The property attributed to the structural messenger of being an unstable intermediate is one of the most specific and novel implications of this scheme; it is required, let us recall, by the kinetics of induction, once the assumption is made that the control systems operate at the genetic level. This leads to a new concept of the mechanism of information transfer, where the protein synthesizing centers (ribosomes) play the role of non-specific constituents which can synthesize different proteins, according to specific instructions which they receive from the genes through M-RNA. The already fairly impressive body of evidence, kinetic and analytical, which supports

this new interpretation of information transfer, is of great interest in itself, even if some of the other assumptions included in the scheme turn out to be incorrect.

These conclusions apply strictly to the bacterial systems from which they were derived; but the fact that adaptive enzyme systems of both types (inducible and repressible) and phage systems appear to obey the same fundamental mechanisms of control, involving the same essential elements, argues strongly for the generality of what may be called "repressive genetic regulation" of protein synthesis.

One is led to wonder whether all or most structural genes (i.e. the synthesis of most proteins) are submitted to repressive regulation. In bacteria, virtually all the enzyme systems which have been adequately studied have proved sensitive to inductive or repressive effects. The old idea that such effects are characteristic only of "non-essential" enzymes is certainly incorrect (although, of course, these effects can be detected only under conditions, natural or artificial, such that the system under study is at least partially non-essential (gratuitous). The results of mutations which abolish the control (such as constitutive mutations) illustrate its physiological importance. Constitutive mutants of the lactose system synthesize 6 to 7% of all their proteins as β-galactosidase. In constitutive mutants of the phosphatase system, 5 to 6% of the total protein is phosphatase. Similar figures have been obtained with other constitutive mutants. It is clear that the cells could not survive the breakdown of more than two or three of the control systems which keep in pace the synthesis of enzyme proteins.

The occurrence of inductive and repressive effects in tissues of higher organisms has been observed in many instances, although it has not proved possible so far to analyse any of these systems in detail (the main difficulty being the creation of controlled conditions of gratuity). It has repeatedly been pointed out that enzymatic adaptation, as studied in micro-organisms, offers a valuable model for the interpretation of biochemical co-ordination within tissues and between organs in higher organisms. The demonstration that adaptive effects in micro-organisms are primarily negative (repressive), that they are controlled by functionally specialized genes and operate at the genetic level, would seem greatly to widen the possibilities of interpretation. The fundamental problem of chemical physiology and of embryology is to understand why tissue cells do not all express, all the time, all the potentialities inherent in their genome. The survival of the organism requires that many, and, in some tissues most, of these potentialities be unexpressed, that is to say *repressed*. Malignancy is adequately described as a breakdown of one or several growth controlling systems, and the genetic origin of this breakdown can hardly be doubted.

According to the strictly structural concept, the genome is considered as a mosaic of independent molecular blue-prints for the building of individual cellular constituents. In the execution of these plans, however, co-ordination is evidently of absolute survival value. The discovery of regulator and operator genes, and of repressive regulation of the activity of structural genes, reveals that the genome contains not only a series of blue-prints, but a co-ordinated program of protein synthesis and the means of controlling its execution.

REFERENCES

Adelberg, E. A. & Umbarger, H. E. (1953). *J. Biol. Chem.* **205**, 475.
Ames, B. N. & Garry, B. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1453.
Ames, B. N., Garry, B. & Herzenberg, L. A. (1960). *J. Gen. Microbiol.* **22**, 369.
Benzer, S. (1953). *Biochim. biophys. Acta*, **11**, 383.
Bertani, G. (1953). *Cold. Spr. Harb. Symp. Quant. Biol.* **18**, 65.

Bertani, G. (1958). *Advanc. Virus Res.* **5**, 151.
Brenner, S., Jacob, F. & Meselson, M. (1961). *Nature*, **190**, 576.
Bussard, A., Naono, S., Gros, F. & Monod, J. (1960). *C. R. Acad. Sci., Paris*, **250**, 4049.
Buttin, G. (1956). Diplôme Et. Sup., Paris.
Buttin, G. (1961). *C. R. Acad. Sci., Paris*, in the press.
Cohen, G. N. & Jacob, F. (1959). *C. R. Acad. Sci., Paris*, **248**, 3490.
Cohen, G. N. & Monod, J. (1957). *Bact. Rev.* **21**, 169.
Cohen, S. S. (1949). *Bact. Rev.* **13**, 1.
Cohen-Bazire, G. & Jolit, M. (1953). *Ann. Inst. Pasteur*, **84**, 1.
Cohn, M. (1957). *Bact. Rev.* **21**, 140.
Cohn, M., Cohen, G. N. & Monod, J. (1953). *C. R. Acad. Sci., Paris*, **236**, 746.
Cohn, M. & Horibata, K. (1959). *J. Bact.* **78**, 624.
Cohn, M. & Monod, J. (1953). In *Adaptation in Micro-organisms*, p. 132. Cambridge University Press.
Cohn, M. & Torriani, A. M. (1952). *J. Immunol.* **69**, 471.
Davern, C. I. & Meselson, M. (1960). *J. Mol. Biol.* **2**, 153.
Demerec, M. (1956). *Cold Spr. Harb. Symp. Quant. Biol.* **21**, 113.
Dienert, F. (1900). *Ann. Inst. Pasteur*, **14**, 139.
Duclaux, E. (1899). *Traité de Microbiologie*. Paris: Masson et Cie.
Echols, H., Garen, A., Garen, S. & Torriani, A. M. (1961). *J. Mol. Biol.*, in the press.
Flaks, J. G. & Cohen, S. S. (1959). *J. Biol. Chem.* **234**, 1501.
Gale, E. F. (1943). *Bact. Rev.* **7**, 139.
Giles, N. H. (1958). *Proc. Xth Intern. Cong. Genetics*, Montreal, **1**, 261.
Gorini, L. & Maas, W. K. (1957). *Biochim. biophys. Acta*, **25**, 208.
Gorini, L. & Maas, W. K. (1958). In *The Chemical Basis of Development*, p. 469. Baltimore: Johns Hopkins Press.
Gros, F., Hiatt, H., Gilbert, W., Kurland, C. G., Risebrough, R. W. & Watson, J. D. (1961). *Nature*, **190**, 581.
Halvorson, H. O. (1960). *Advanc. Enzymol.* in the press.
Hartman, P. E., Loper, J. C. & Serman, D. (1960). *J. Gen. Microbiol.* **22**, 323.
Herzenberg, L. (1959). *Biochim. biophys. Acta*, **31**, 525.
Hogness, D. S., Cohn, M. & Monod, J. (1955). *Biochim. biophys. Acta*, **16**, 99.
Jacob, F. (1954). *Les Bactéries Lysogènes et la Notion de Provirus*. Paris: Masson et Cie.
Jacob, F. (1960). *Harvey Lectures*, 1958–1959, series **54**, 1.
Jacob, F. & Adelberg, E. A. (1959). *C.R. Acad. Sci., Paris*, **249**, 189.
Jacob, F. & Campbell, A. (1959). *C.R. Acad. Sci., Paris*, **248**, 3219.
Jacob, F., Fuerst, C. R. & Wollman, E. L. (1957). *Ann. Inst. Pasteur*, **93**, 724.
Jacob, F. & Monod, J. (1959). *C.R. Acad. Sci., Paris*, **249**, 1282.
Jacob, F., Perrin, D., Sanchez, C. & Monod, J. (1960*a*). *C.R. Acad. Sci., Paris*, **250**, 1727.
Jacob, F., Schaeffer, P. & Wollman, E. L. (1960*b*). In *Microbial Genetics*, Xth Symposium of the Society for General Microbiology, p. 67.
Jacob, F. & Wollman, E. L. (1953). *Cold Spr. Harb. Symp. Quant. Biol.* **18**, 101.
Jacob, F. & Wollman, E. L. (1956). *Ann. Inst. Pasteur*, **91**, 486.
Jacob, F. & Wollman, E. L. (1957). In *The Chemical Basis of Heredity*, p. 468. Baltimore: Johns Hopkins Press.
Kaiser, A. D. (1957). *Virology*, **3**, 42.
Kaiser, A. D. & Jacob, F. (1957). *Virology*, **4**, 509.
Kalckar, H. M., Kurahashi, K. & Jordan, E. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1776.
Karstrom, H. (1938). *Ergebn. Enzymforsch.* **7**, 350.
Képès, A. (1960). *Biochim. biophys. Acta*, **40**, 70.
Képès, A., Monod, J. & Jacob, F. (1961). In preparation.
Kogut, M., Pollock, M. & Tridgell, E. J. (1956). *Biochem. J.* **62**, 391.
Kornberg, A., Zimmerman, S. B., Kornberg, S. R. & Josse, J. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 772.
Kruh, J., Rosa, J., Dreyfus, J.-C. & Schapira, G. (1961). *Biochim. biophys. Acta*, in the press.
Lamfrom, H. (1961). *J. Mol. Biol.* **3**, 241.
Lederberg, E. (1960). In *Microbial Genetics*, The Xth Symposium of the Society of General Microbiology, p. 115.

Levinthal, C. (1959). In *Structure and Function of Genetic Elements*, Brookhaven Symposia in Biology, p. 76.
Luria, S. E. & Human, M. L. (1950). *J. Bact.* **59**, 551.
Lwoff, A. (1953). *Bact. Rev.* **17**, 269.
Lwoff, A., Siminovitch, L. & Kjeldgaard, N. (1950). *Ann. Inst. Pasteur*, **79**, 815.
Magasanik, B., Magasanik, A. K. & Neidhardt, F. C. (1959). In *A Ciba Symposium on the Regulation of Cell Metabolism*, p. 334. London: Churchill.
Monod, J. (1942). *Recherches sur la Croissance des Cultures Bactériennes*. Paris: Hermann.
Monod, J. (1955). *Exp. Ann. Biochim. Méd.* série XVII, p. 195. Paris: Masson et Cie.
Monod, J. (1956). In *Units of Biological Structure and Function*, p. 7. New York: Academic Press.
Monod, J. (1958). *Rec. Trav. Chim. des Pays-Bas*, **77**, 569.
Monod, J. (1959). *Angew. Chem.* **71**, 685.
Monod, J. & Audureau, A. (1946). *Ann. Inst. Pasteur*, **72**, 868.
Monod, J. & Cohen-Bazire, G. (1953). *C.R. Acad. Sci., Paris*, **236**, 530.
Monod, J. & Cohn, M. (1952). *Advanc. Enzymol.* **13**, 67.
Monod, J. & Cohn, M. (1953). In *Symposium on Microbial Metabolism*. VIth Intern. Cong. of Microbiol., Rome, p. 42.
Monod, J., Pappenheimer, A. M. & Cohen-Bazire, G. (1952), *Biochim. biophys. Acta*, **9**, 648.
Monod, J. & Wollman, E. L. (1947). *Ann. Inst. Pasteur*, **73**, 937.
Naono, S. & Gros, F. (1960*a*). *C.R. Acad. Sci., Paris*, **250**, 3527.
Naono, S. & Gros, F. (1960*b*). *C.R. Acad. Sci., Paris*, **250**, 3889.
Neidhardt, F. C. & Magasanik, B. (1956*a*). *Nature*, **178**, 801.
Neidhardt, F. C. & Magasanik, B. (1956*b*). *Biochim. biophys. Acta*, **21**, 324.
Novick, A. & Szilard, L., in Novick, A. (1955). *Ann. Rev. Microbiol.* **9**, 97.
Pardee, A. B. (1957). *J. Bact.* **73**, 376.
Pardee, A. B., Jacob, F. & Monod, J. (1959). *J. Mol. Biol.* **1**, 165.
Pardee, A. B. & Prestidge, L. S. (1959). *Biochim. biophys. Acta*, **36**, 545.
Pardee, A. B. & Prestidge, L. S. (1961). In preparation.
Perrin, D., Bussard, A. & Monod, J. (1959). *C.R. Acad. Sci., Paris*, **249**, 778.
Perrin, D., Jacob, F. & Monod, J. (1960). *C.R. Acad. Sci., Paris*, **250**, 155.
Pollock, M. (1950). *Brit. J. Exp. Pathol.* **4**, 739.
Pollock, M. & Perret, J. C. (1951). *Brit. J. Exp. Pathol.* **5**, 387.
Pontecorvo, G. (1958). *Trends in Genetic Analysis*. New York: Columbia University Press.
Rickenberg, H. V., Cohen, G. N., Buttin, G. & Monod, J. (1956). *Ann. Inst. Pasteur*, **91**, 829.
Riley, M., Pardee, A. B., Jacob, F. & Monod, J. (1960). *J. Mol. Biol.* **2**, 216.
Rotman, B. & Spiegelman, S. (1954). *J. Bact.* **68**, 419.
Siminovitch, L. & Jacob, F. (1952). *Ann. Inst. Pasteur*, **83**, 745.
Stanier, R. Y. (1951). *Ann. Rev. Microbiol.* **5**, 35.
Szilard, L. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 277.
Torriani, A. M. (1960). *Biochim. biophys. Acta*, **38**, 460.
Umbarger, H. E. (1956). *Science*, **123**, 848.
Vogel, H. J. (1957*a*). *Proc. Nat. Acad. Sci., Wash.* **43**, 491.
Vogel, H. J. (1957*b*). In *The Chemical Basis of Heredity*, p. 276. Baltimore: Johns Hopkins Press.
Volkin, E. & Astrachan, L. (1957). In *The Chemical Basis of Heredity*, p. 686. Baltimore: Johns Hopkins Press.
Went, F. C. (1901). *J. Wiss. Bot.* **36**, 611.
Wijesundera, S. & Woods, D. D. (1953). *Biochem. J.* **55**, viii.
Willson, C., Perrin, D., Jacob, F. & Monod, J. (1961). In preparation.
Wollman, E. L. & Jacob, F. (1959). *La Sexualité des Bactéries*. Paris: Masson et Cie.
Yanofsky, C. (1960). *Bact. Rev.* **24**, 221.
Yanofsky, C. & Lennox, E. S. (1959). *Virology*, **8**, 425.
Yarmolinsky, M. B. & Wiesmeyer, H. (1960). *Proc. Nat. Acad. Sci., Wash.* in the press.
Yates, R. A. & Pardee, A. B. (1956). *J. Biol. Chem.* **221**, 757.
Yates, R. A. & Pardee, A. B. (1957). *J. Biol. Chem.* **227**, 677.
Yčas, M. & Vincent, W. S. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 804.
Zabin, I., Képès, A. & Monod, J. (1959). *Biochem. Biophys. Res. Comm.* **1**, 289.

J. Mol. Biol. (1961) **3**, 357–361

Letters to the Editor

Temperature Dependence of the Adsorption of Native DNA in a Polarized Water-Mercury Interface

In a previous communication (Miller, 1961) data were given for the areas per nucleotide of *E. coli* RNA, denatured herring sperm DNA and native salmon sperm DNA adsorbed on mercury-water interfaces at 25 °C. The areas were obtained by extrapolation of the surface concentration Γ_t calculated by equation (1) (Koryta, 1953; Delahay & Trachtenberg, 1957)

$$\Gamma_t = 0{\cdot}745\, D^{\frac{1}{2}} t^{\frac{1}{2}} C_b \qquad (1)$$

toward nucleic acid concentrations C_b and drop ages t at which the surface just becomes covered. Equation (1) is valid provided the adsorption on the surface of a growing mercury droplet is diffusion controlled, the adsorption process being instantaneous when the nucleic acid reaches the surface (Miller, 1961).

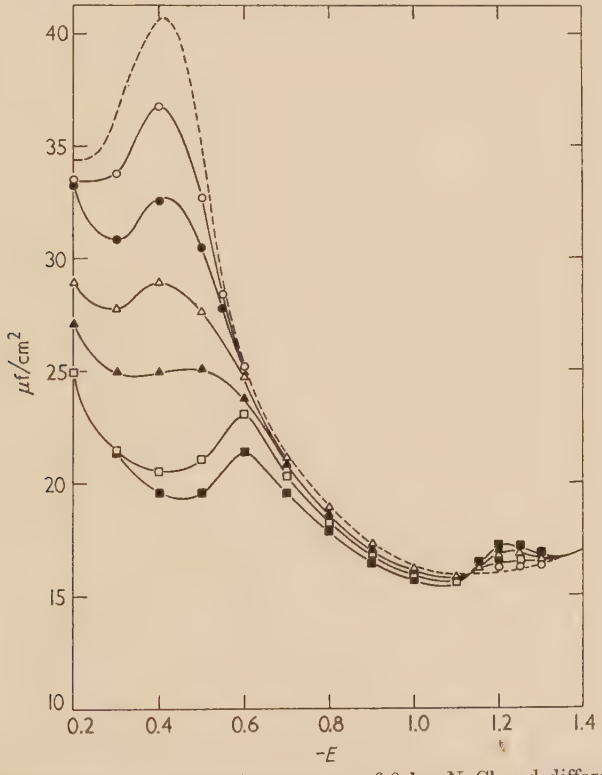


Fig. 1. Differential capacity curves in the presence of 0·1 N-NaCl and different concentrations of salmon sperm DNA at 0°C:
- - - — no DNA; ○ — C_b = 0·04 mg/ml.; ● — C_b = 0·08 mg/ml.; △ — C_b = 0·12; ▲ — C_b = 0·16 mg/ml.; □ — C_b = 0·28 mg/ml.; ■ — C_b = 0·6 mg/ml.

The present communication is concerned with variations in adsorption of native salmon sperm DNA with changes in temperature as revealed by differential capacity measurements. Measurements of the differential capacity of the double layer on a

polarized mercury surface (Grahame, 1949) were carried out at 0 °C, 15 °C, 25 °C, 30 °C and 35 °C. The experimental conditions were similar to those reported in the previous communication (Miller, 1961). A dropping mercury electrode was used. It was polarized relative to a 1 N-calomel electrode in the presence of 0·1 N-NaCl as supporting electrolyte at various concentrations of DNA at pH 6. The drop age at the time of the measurement varied between 5 and 20 sec.

There is in general a variation of the differential capacity curves with temperature. One set of these curves for 0 °C is given in Fig. 1. The temperature effects are greatest at potentials where the mercury surface is negatively charged, namely at potentials more negative than — 0·6 v relative to the 1 N-calomel electrode. In this region the limiting surface area per nucleotide, as reported previously (Miller, 1961) for the fully

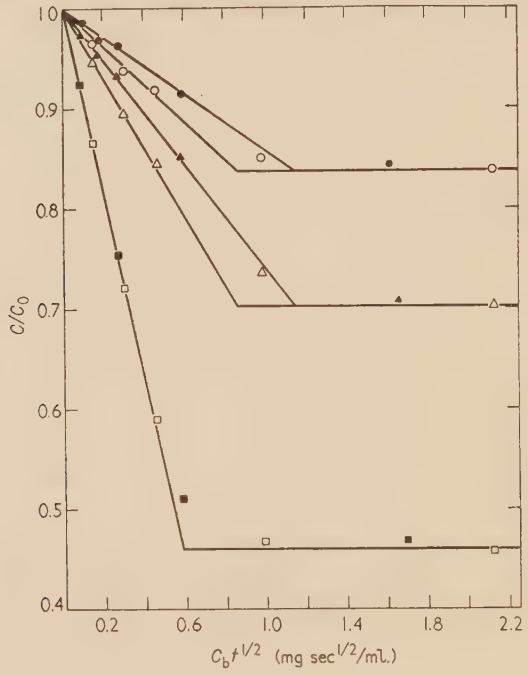


FIG. 2. Reduced capacity C/C_0 *versus* concentration times the square root of the drop age $C_b t^{\frac{1}{2}}$;

— E relative to N-calomel electrode	t = 14 sec	t = 5 sec
0·8	○	●
0·6	△	▲
0·4	□	■

covered surface at 25 °C, is about 35 Å². The limiting surface areas per nucleotide obtained by the same method at different temperatures between 0 ° to 30 °C are the same within experimental error. The area at these temperatures is independent of the age of the drop between 5 and 20 sec. At 35 °C the limiting area is about 45 Å²/nucleotide at 5 sec and increases at higher drop ages (Fig. 2). This increase in area is attributed to an unfolding of the DNA molecules in the surface. The measured area σ is thought to be made up according to equation (2)

$$\sigma = \sigma_{nat} + \alpha(\sigma_{unf} - \sigma_{nat}) \qquad (2)$$

where σ_{nat} is the area per nucleotide of a native DNA molecule, σ_{unf} is the area per nucleotide of an unfolded molecule and α is the fraction unfolded at the time of

the measurement. As σ_{nat} is about 35 Å²/nucleotide and σ_{unf} is about 90 Å²/nucleotide, the area of 45 Å²/nucleotide at 5 sec and of 60 Å²/ nucleotide at 14 sec indicates that about 15% and 38% respectively of the molecules has become unfolded.

At temperatures lower than 35 °C no unfolding process on the negatively charged surface is noticed. But even here the temperature influences the effect of native DNA on the differential capacity of the electrical double layer. In Fig. 3 the logarithm of

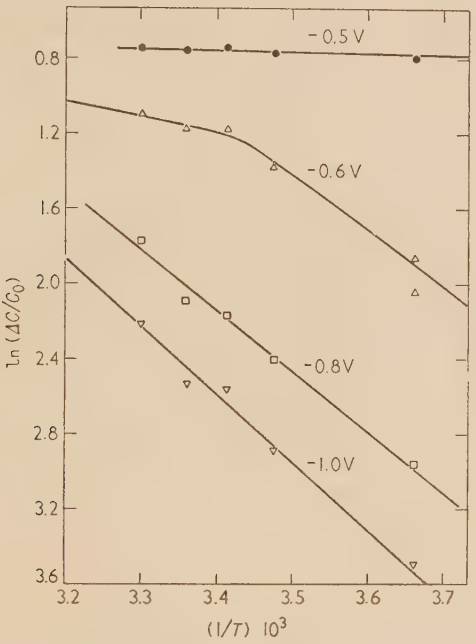


FIG. 3. $\ln(\Delta C/C_0)$ *versus* the reciprocal of the absolute temperature for different potentials.

the lowering of differential capacity divided by the capacity obtained in the presence of the supporting electrolyte only, $\ln(\Delta C/C_0)$, was plotted against the reciprocal of the absolute temperature. In order to assure that measurements at equal surface concentration were compared, ΔC values corresponding to constant $D^{\frac{1}{2}}t^{\frac{1}{2}}C_b$ were taken at all temperatures. Straight lines were obtained with slopes which depend on the potential. At potentials more positive than — 0·6 V with respect to the N-calomel electrode (corresponding approximately to zero charge) the lines are almost parallel to the axis. However at more negative potentials the slope of the lines indicates a considerable temperature dependence of the influence of the adsorbed DNA on the differential capacity. In order to explain the described behavior the following model is proposed.

At temperatures below 35 °C the double-helical structure of the native DNA is preserved, provided that the mercury surface is not positively charged and does not destroy the helical structure by long range electrostatic forces. There is, however, always a certain probability of breaking the pair of hydrogen bonds which hold together two nucleotides. If this happens the nucleotide can rotate around the sugar-phosphate bonds. In general at room temperature the recombination rate is considerably larger than the cleavage and only a small fraction of the nucleotides is in the uncoupled state. On the surface, however, a free nucleotide rotating around the sugar-phosphate bond

can hit the surface and adhere. Consequently the uncoupling tendency of the nucleotide pairs is enhanced by the nearness of the surface. The path of the rotating nucleotide is determined by its position in the helix and, provided the position of the DNA molecule in the surface is fixed, the rotating nucleotide can assume only a limited number of positions in the surface. For the same reason the competition for the same site in the surface by different nucleotides will be limited as well.

As a first approximation we assume that a nucleotide pair can exist either coupled together in the double helix or dissociated and adhering to the surface by the purine or pyrimidine residue. For this adhesion a nucleotide has at its disposal a set of p sites in the surface. A further assumption now is that two residues cannot occupy the same set of sites and that these sets of sites are also not available to other nucleotides. In this case

$$X/(1 - X) = p^2 \, e^{-(\Delta H_{unf} + \Delta H_{ads})/RT} = p^2 \, e^{-\Delta H/RT} \tag{3}$$

where X is the fraction of residues adhering to the surface, ΔH_{unf} is the difference in enthalpy/residue in going from the coupled to the unfolded state and ΔH_{ads} is the difference in enthalpy/residue in going from the unfolded to the adsorbed state.

By differentiating equation (2) with respect to $1/T$ we obtain

$$d\ln X d(1/T) = - (\Delta H/R)/(1 + p^2 \, e^{-\Delta H/RT}) \tag{4}$$

In the event that the same set of sites can be occupied by different residues, the denominator of the right side of equation (4) is changed. For instance, if the sites in the domain of one residue could also be occupied by its two closest neighbours on the same strand, half of the sites being in reach of each neighbour, equation (4) would have the following form

$$d\ln X/d(1/T) = - (\Delta H/R)/[1 + (2p + p^2)e^{-\Delta H/RT}] \tag{5}$$

It is assumed that the lowering of the capacity ΔC by adsorbed DNA is proportional to the number of residues anchored in the surface (Melik-Gaikazyan, 1952). Only the anchored residues affect the differential capacity (Miller & Grahame, 1956) by lowering the dielectric constant at the mercury surface and by increasing the thickness of the electrical double layer. Assuming this, equation (3) can be written

$$d\ln(\Delta C/C_0)/d(1/T) = - (\Delta H/R)/(1 + p^2 \, e^{-\Delta H/RT}) \tag{6}$$

where Δ was divided by C_0 in order to allow for the variation of the surface capacity with temperature due to those factors which are independent of the adsorption.

The plot of $\ln(\Delta C/C_0)$ against $1/T$ in Fig. 3 gives straight lines at adequately negative potentials. According to equation (5) the slope of the lines is equal to or smaller than $\Delta H/R$. At $-$ 1·0 v with respect to the N-calomel electrode the slope is about 3500 °K. Assuming that the denominator on the right hand side of equation (4) or (6) is 1 we find a value for $\Delta H/RT$ of about 12. Substituting this value in the expression for the denominator on the right hand side of equation (4) or (6) we find that our assumption that it is close to 1 was justified even if p^2 were extremely large (say 1000). Accordingly the slopes of the straight lines in Fig. 3 at negative potentials give the value of $\Delta H/R$. At about $-$ 1·3 v relative to the N-calomel electrode, which is the desorption potential, $\Delta H_{ads} = 0$ and $\Delta H = \Delta H_{unf}$. The slope extrapolated at this potential gives the value for ΔH_{unf} of about 7·5 to 8 kcal. This value is close to but somewhat lower than the ΔH_{unf} value of 9 kcal obtained by Warner & Breslow (1958) for 1 : 1 complexes of polyadenylic and polyuridylic acid.

The critical reading of the manuscript by Dr. A. Silberberg and Professor P. Doty's gift of the salmon sperm DNA are thankfully acknowledged.

Polymer Department
Weizmann Institute of Science
Rehovot
Israel

I. R. Miller

Received 23 August 1960

REFERENCES

Delahay, P. & Trachtenberg, I. (1957). *J. Amer. Chem. Soc.* **79**, 2235.
Grahame, D. C. (1949). *J. Amer. Chem. Soc.* **71**, 2975.
Koryta, J. (1953). *Collect. Czech. Chem. Com.* **18**, 206.
Melik-Gaikazyan, V. I. (1952). *J. Phys. Chem. U.S.S.R.* **26**, 60.
Miller, I. R. (1961). *J. Mol. Biol.* **3**, 229.
Miller, I. R. & Grahame, D. C. (1956). *J. Amer. Chem. Soc.* **78**, 3577.
Warner, R. C. & Breslow, E. (1958). *Proc. 4th Int. Congr. Biochem., Vienna*, p. 157.

J. Mol. Biol. (1961) **3**, 362–366

Comparative Studies on Tryptic Digests of Myoglobins

The success of the fingerprint method of Ingram (1958) in demonstrating small chemical differences between molecular species of human haemoglobins suggested its use in comparing myoglobins of various animal species. Myoglobins obtained from a number of species have been compared both crystallographically and immunologically, but there is little information available on the chemical differences between species (Kendrew, Parrish, Marrack & Orlans, 1954). It might be expected that fingerprinting would show differences of only a few peptides in closely related species and larger differences as the relationship became more distant. Indeed this type of relationship has been observed in a recent comparison of animal haemoglobins by the fingerprinting technique (Zuckerkandl, Jones & Pauling, 1960). It was surprising, therefore, to discover that the fingerprint patterns of myoglobins showed certain similarities among all the species studied but, in general, the differences between closely related species were just as marked as those between very different types of animal.

The myoglobins used in this study had been prepared by Kendrew and others (Kendrew *et al.*, 1954) and stored as an ammonium sulphate paste at 3°C. Each was dialysed against 10^{-3} M-phosphate buffer at pH 7 and then submitted to tryptic digestion and fingerprinting (Ingram, 1958). For each myoglobin approximately 23 discrete spots which gave a purple colour with ninhydrin were observed. It has been shown by Edmundson & Hirs (1959) that 20 soluble peptides which are obtained on tryptic digestion of sperm whale myoglobin account for approximately 65% of the protein, the remaining "core" being resistant to digestion by either trypsin or chymotrypsin.

Specific tests for histidine, arginine, tyrosine and sulphur-containing amino acids were carried out on duplicate fingerprints of each species.† Most of the peptides occurring in the fingerprint of each myoglobin were also found in one or more other species of myoglobin. The criteria for assuming that two peptides were similar were that they appeared in identical positions on the fingerprint and that they gave identical reactions to the tests for specific amino acids. This assumption is subject to certain limitations as discussed below and by Ingram (1958) and Zuckerkandl *et al.* (1960). Two representative fingerprints are shown in Fig. 1. Table 1 gives all of the peptides found in at least two species of myoglobin (the numbered peptides of Fig. 1). The fingerprints of each species of myoglobin also contain 5 to 7 unique peptides which are not observed in the fingerprint of any other species studied. Each unique tryptic peptide of myoglobin must represent at least one difference in amino acid sequence between that species and any other. Thus there is a minimum of 5 chemical differences between any two species of the present study, even those most closely related. In the studies of Kendrew *et al.* (1954), sperm whale myoglobin was found to resemble the myoglobins from porpoise and sei whale both crystallographically and immunologically. There was less immunological cross-reaction between sperm whale

† The colour reactions specific for various amino acids were carried out as described by Block, Durrum & Zweig (1958) under the following authors: ninhydrin, Toennies & Kolb (1951); sulphur, Toennies & Kolb (1951); histidine, Karler (1952) and Baldridge & Lewis (1953); arginine, Jepson & Smith (1953); tyrosine, Acher & Crocker (1952); tryptophan, Smith (1953).

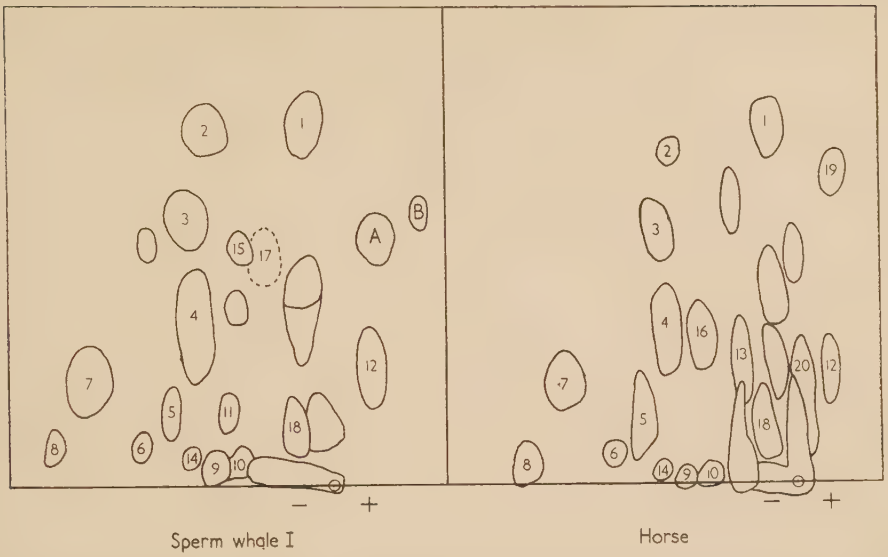


FIG. 1. Tracings of fingerprints of tryptic digests of fraction I of sperm whale myoglobin and of horse myoglobin. Peptide 17 does not occur in sperm whale myoglobin, but is inserted in dotted lines to show its relative position on the fingerprint.

TABLE 1

Peptides found in two or more species of myoglobin

Peptide no.	Positive tests for specific amino acids	Species						
		sperm whale	horse	seal	porpoise	sei whale	gentoo penguin	tortoise
1	arg	+	+	+	+	+	+	+
2	none	+	+	+	+	+	+	+
3	none	+	+	+	+	+	0	+
4	his	+	+	+	+	+	0	+
5	his	+	+	+	+	+	+	+
6	his	+	+	+	+	+	+	+
7	none	+	+	+	+	+	+	+
8	none	+	+	+	+	+	0	0
9	his	+	+	+	+	+	+	0
10	his	+	+	+	+	+	+	+
11	none	+	0	0	+	+	0	0
12	sulphur	+	+	0	+	+	+	0
13	his	0	+	+	0	+	0	0
14	his	+	+	0	0	0	0	0
15	arg	+	0	0	+	0	0	+
16	none	0	+	+	+	+	0	0
17	his	0	0	+	+	+	+	0
18	his, sulphur	+	+	0	+	+	+	+
19	none	0	+	0	0	+	+	+
20	his	0	+	0	0	+	0	0

Presence of peptide, +. Absence of peptide, 0.
The specific tests were for histidine, tyrosine, arginine and sulphur-containing amino acids.

and seal myoglobins and no cross-reaction between the myoglobins of sperm whale and horse. It was thus surprising that a similar relationship was not observed when the fingerprints were compared. The myoglobins of sperm whale and porpoise possess 14 peptides in common, and sperm whale and sei whale have 13 in common. As might be expected from the immunological findings, sperm whale and seal myoglobins are less similar and have only 10 peptides in common. On the other hand, sperm whale and horse myoglobin would appear to resemble each other much more closely than the immunological and crystallographical data would predict, having 13 peptides in common.

This method is, of course, subject to certain limitations. One cannot assume identity of 2 peptides which appear in identical positions on a fingerprint, for small differences in amino acid composition or sequence might not be detectable. This is particularly true for the larger peptides. Thus one can use data of this type to determine a minimum number of differences that must exist between two proteins, but they may, in fact, be more numerous. In addition 40% of the protein (in the case of sperm whale myoglobin) is not digested by trypsin, so that many other chemical similarities and differences may not be observed in the present study. These might include those differences responsible for immunological specificity.

It is known that the sperm whale myoglobin of these preparations is not homogeneous, but can be separated into 4 or possibly 5 components on Amberlite IRC-50 (Edmundson & Hirs, 1959). In this laboratory sperm whale myoglobin has been separated into two components on a carboxymethylcellulose column equilibrated with sodium phosphate buffer, pH 7·2, which was 0·02 M with respect to sodium. Fraction I is eluted with the latter buffer and fraction II, which is the major portion of the protein, is eluted by increasing the sodium concentration to 0·1 M (L. K. Steinrauf, unpublished data). These two fractions were each subjected to tryptic digestion and fingerprinting. The fingerprints were found to be identical except for the presence of one additional negatively-charged tyrosine-containing peptide in fraction I (peptide B of Fig. 1). Differences involving only one peptide are also obtained when comparing most of the various molecular species of adult human haemoglobins (Ingram, 1958) and adult horse haemoglobins (Perutz, Steinrauf, Stockell & Bangham, 1959). It thus appears that the chemical differences between like proteins of a given species are very much less than between different, but closely related, species.

The chemical differences between fractions I and II of sperm whale myoglobin do not appear to be due to substitution of one amino acid in fraction I by another in fraction II, but to the presence of one additional peptide in fraction I. Since peptide A also gave a positive test for tyrosine, peptides A and B were eluted and subjected to hydrolysis and amino acid analysis (Redfield, 1953, as modified by Hunt & Ingram, 1960). The peptides were each found to contain glutamic acid, glycine, leucine and tyrosine. As no lysine or arginine is present in these tryptic peptides, it is assumed that they represent carboxyl-terminal peptides of the proteins. This is in accord with Edmundson's finding (1960) that the carboxyl-terminal tryptic peptide of myoglobin has the sequence glu.leu.gly.tyr.glu-NH_2. It is at present difficult to explain why myoglobin I contains two such peptides. It is possible that it is a mixture of two molecular species, the "components II and III" reported by Edmundson (1960). The chemical differences between peptides A and B could be explained by the presence of glutamic acid rather than glutamine in peptide B. This would account for the greater net negative charge of peptide B. However, proof of this must await more intensive study of the chemistry of peptide B.

Scouloudi (1959) has shown that seal myoglobin, though crystallizing in a different form from sperm whale myoglobin, has a very similar three-dimensional structure. In their determination of the structure of haemoglobin, Perutz *et al.* (1960) have shown that the three-dimensional coiling of each of the polypeptide chains of horse haemoglobin is similar to that of sperm whale myoglobin. Thus it was of interest to compare fingerprints of tryptic digests of haemoglobin and myoglobin from the same species (Fig. 1 and Fig. 2). It might be expected that a few peptides in common would be

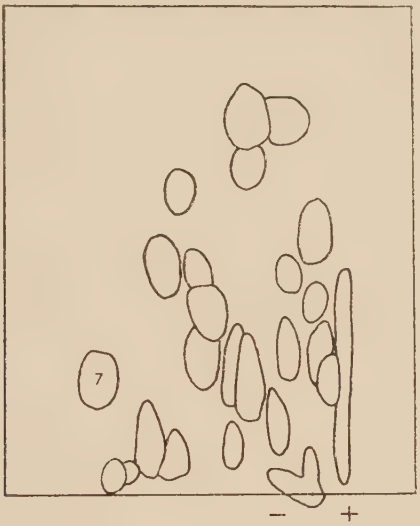


Fig. 2. Tracings of the fingerprint of a tryptic digest of horse haemoglobin.

observed, but only peptide 7 appeared to be the same in the two proteins obtained from horse. Elution of this peptide from fingerprints of both proteins and determination of its amino acid composition showed that lysine was the only amino acid present. It thus follows that large chemical differences can exist between proteins without altering the specific folding of their polypeptide chains. Zuckerkandl *et al.* (1960) reported the presence of free lysine in tryptic digests of all the haemoglobins investigated and suggested that the sequence lys.lys. or arg.lys. is common to all animal haemoglobins. It appears that this generalization applies to myoglobins as well, since free lysine was observed in tryptic digests of all of the myoglobins included in the present study. The fingerprint method is thus useful in studies of similar proteins obtained from different biological species, as similarities between proteins which do not cross-react immunologically may be detected by this method.

I am indebted to Dr. J. C. Kendrew for suggesting this study and for providing the myoglobin preparations investigated. The capable technical assistance of Miss Eileen Blincoe is greatly appreciated. This investigation has been aided by grants from the U.S. Public Health Service and the Jane Coffin Childs Memorial Fund for Medical Research.

Anne Stockell†

Medical Research Council Unit
for Molecular Biology
Cavendish Laboratory
Cambridge, England

Received 23 December 1960

† U.S. Public Health Service Research Fellow of the National Institute of Arthritis and Metabolic Diseases 1958–1959. Fellow of the Jane Coffin Childs Memorial Fund for Medical Research 1959–1960. Now at Department of Biochemistry, University of Cambridge, Cambridge, England.

REFERENCES

Acher, R. & Crocker, C. (1952). *Biochim. biophys. Acta*, **9**, 704.
Baldridge, R. C. & Lewis, H. B. (1953). *J. Biol. Chem.* **202**, 169.
Block, R. J., Durrum, E. L. & Zweig, G. (1958). *Paper Chromatography and Paper Electrophoresis*. New York: Academic Press.
Edmundson, A. B. (1960). Ph.D. Thesis, The Rockefeller Institute.
Edmundson, A. B. & Hirs, C. H. W. (1959). *Fed. Proc.* **18**, 220
Hunt, J. A. & Ingram, V. M. (1960). *Biochim. biophys. Acta*, **42**, 409.
Ingram, V. M. (1958). *Biochim. biophys. Acta*, **28**, 539.
Jepson, J. B. & Smith, I. (1953). *Nature*, **172**, 1100.
Karler, A. (1952). *U.S. Atomic Energy Comm.* URCL-1838, May 29.
Kendrew, J. C., Parrish, R. G., Marrack, J. R. & Orlans, E. S. (1954). *Nature*, **174**, 946.
Perutz, M. F., Steinrauf, L. K., Stockell, A. & Bangham, A. D. (1959). *J. Mol. Biol.* **1**, 402.
Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. C. T. (1960). *Nature*, **185**, 416.
Redfield, R. R. (1953). *Biochim. biophys. Acta*, **10**, 344.
Scouloudi, H. (1959). *Nature*, **183**, 374.
Smith, I. (1953). *Nature*, **171**, 43.
Toennies, G. & Kolb, J. J. (1951). *Analyt. Chem.* **23**, 823.
Zuckerkandl, E., Jones, R. T. & Pauling, L. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 1349.

INSTRUCTIONS TO AUTHORS

1. The *Journal of Molecular Biology* will publish papers on the nature, production and replication of biological structure at the molecular level, and its relation to function. Suitable subjects are sub-cellular organization: molecular genetics: structure and replication of viruses: molecular structure of muscle, nerve and other tissues: structure of proteins, nucleic acids, carbohydrates, lipids, etc., and their synthetic analogues, as investigated by X-rays, light absorption and other methods: problems of inter- and intra-molecular energy transfer. The Journal will not as a rule publish papers when the subject matter is outside the fields defined above; however, when the subject matter borders on molecular biology and is of special interest in relation to molecular biology such papers may be published when space permits.

2. The Journal will also publish Letters to the Editor (averaging 1000 words), and occasional review articles dealing with a more restricted or immediately topical field than those in review journals. (The Editor should be consulted in advance by anyone who considers submitting such a review.)

3. Communications will often describe results obtained by specialized techniques (for example, those of X-ray crystallography and of genetics), and it is most important that technical terms should be avoided or be clearly defined; in any event the actual results obtained and their implications should be described in a way which does not demand from the reader specialized knowledge of a particular technique.

Papers should contain the maximum density of information compatible with clarity, and every sentence should be scrutinized to ensure that its presence is essential. Authors should include enough experimental details to enable others to repeat the work (these may be printed in small type). Careful preparation of manuscript will aid prompt publication; carelessness will cause delay.

4. *Submission of Manuscripts.* Manuscripts should be sent to Dr. J. C. Kendrew, Peterhouse, Cambridge, England, or to any member of the Editorial Board. The submission of manuscript will be taken to imply that the material is original, and that no similar paper has been or will be submitted for publication elsewhere.

Papers will normally be published in English. Manuscripts in French or German will also be accepted, though the use of these languages may cause delay in publication.

5. *Copy.* Papers should be typewritten in double spacing *throughout* (including tables, legends and footnotes), on one side of quarto paper, with a margin of 1 inch all round. The position of tables and illustrations should be indicated in the text; tables and legends for illustrations should be typed separately at the end of the manuscript. All pages should be numbered serially and securely fastened together. Two copies of each manuscript should be submitted, together with three copies of the summary (see § 6) and four copies of the abstract (see § 9).

A short running title of not more than 40 characters (including spaces), suitable for page headings, should be indicated if the full title is longer than this.

The name of the laboratory where the work has been done should be indicated on the title page. All typescript (including the references) should be carefully checked for errors.

6. *Summary.* A summary will be printed at the head of all papers (excluding Letters to the Editor); this should generally not exceed 300 words, and should be intelligible to the general reader without reference to the main text. Abbreviations should be avoided. Summaries of papers to be printed in French or German will be in English, and should be longer (up to 1000 words) for the benefit of readers not familiar with these languages.

7. *Conventions.* The conventions used in current issues of the Journal for headings, references, etc., should be followed in preparing manuscripts. References should be listed alphabetically at the end of the paper, and abbreviated according to the *World List of Scientific Periodicals*, except that for periodicals with English names capital initial letters should be used throughout. Citations in the text should read thus: Smith, Brown & Robinson (1957), or (Smith, Brown & Robinson, 1957). When a citation has more than two authors, the names of all of them should be given at the first citation (except for more than five authors), and on subsequent citations thus: (Smith *et al.*, 1957). The conventions (Brown, 1957*a*), (Brown, 1957*b*) should be used where more than one paper by the same author(s) has appeared in one year. Volume number should be indicated in the typescript thus: **35.**

Footnotes should be avoided whenever possible. Essential footnotes should be indicated by superscript figures in the text and collected on a single page placed at the end of the manuscript.

Customary abbreviations will be used; authors may refer to *Chemical Abstracts*, to the *Biochemical Journal*, **66**, 8 (1957), and to *Notes on the preparation of papers to be communicated to the Royal Society* (1951). Certain non-standard abbreviations of chemical substances are convenient, and those listed in the *Journal of Biological Chemistry* may be used without further definition (e.g. DNA, ATP). Any other abbreviations of this kind should be used sparingly, and must be defined in a footnote at the point in each manuscript where the first such abbreviation occurs.

8. *Illustrations.* It is often convenient to make a *drawing* two or three times larger than it will appear in the Journal, but very large drawings should be avoided and it will be helpful if drawings and illustrations are submitted on quarto (or 8½ by 11 inch) sheets. Drawings should be made in Indian ink on tracing linen, Bristol board, or faintly blue-lined graph paper; the originals should always be submitted with the manuscript, since photo-prints of drawings often cannot be reproduced satisfactorily. Explanations should, as far as possible, be placed in the legend; the drawing itself should contain a minimum of lettering, and this should be inserted lightly in *blue pencil*, in lower case except for proper names, when a capital should be used. It should be remembered that drawings will be reduced to a size not exceeding 8 by 5 inches, and that the final thickness of the lines should be 0·25 to 0·3 mm. If a curve is given merely for illustrative purposes no cross-lines are needed, but the axes should be graduated at intervals, and the diagram should be boxed in. Where possible, related diagrams should be grouped to form a single figure, and should be prepared uniformly and *on the same scale.* Where appropriate and possible a scale should be included in the illustration itself; the magnification may be noted in the legend and adjusted after reduction by the blockmaker.

Whenever possible *photographs* should be supplied the same size as they are to appear in the Journal. The page area available for a plate is 5 by 7 inches or, at the most, 6 by 8 inches.

Line drawings will be referred to as Figure 1, Figure 2, etc., and half-tone blocks as Plate I, Plate II, etc. The name(s) of the author(s) and the number should be written on all illustrations, together with the indication "Top" if necessary. Besides the original illustrations *one* duplicate set must accompany the second copy of the manuscript; these may be photo-prints, and need not be of the same quality as the originals but must be clear and legible. If *oversize* original drawings are submitted (bigger than 8½ by 11 inches) they must be accompanied by *two* sets of small photo-prints.

9. *Abstracts.* Authors should send with their manuscripts *four copies* (typewritten in double spacing) of an abstract, in about 100 words, suitable for inclusion in biological abstracting journals. This abstract will not appear in the *Journal of Molecular Biology* but will be scrutinized by the Editors before being passed for publication. The abstract should outline as briefly as possible the results and definitive conclusions of the work. Details of methods are generally not required.

The abstract should be headed as follows: title of paper; name(s) of author(s); *J. Mol. Biol.*, followed by space for the year, volume and page references; abbreviated postal address (including country of origin) to which applications for reprints may be sent. For example: The isolation and properties of deoxyribonucleoprotein particles containing single nucleic acid molecules. G. Zubay & P. Doty. J. Mol. Biol. . . . . (Dept. of Chem., Harvard Univ., Cambridge, Mass., U.S.A.). —text of abstract follows.

10. *Reprints.* Fifty reprints of each paper will be supplied free of charge; additional copies may be purchased.

Protein Biosynthesis

A Symposium held at Wassenaar, 29 August to 2 September, 1960, under the auspices of UNESCO and the Council for International Organizations of Medical Sciences

Edited by R. J. C. HARRIS, *Imperial Cancer Research Fund, London*

April 1961, 409 pages, 95s.

THIS INTERNATIONAL SYMPOSIUM was sponsored jointly by UNESCO, CIOMS, and the Netherlands Cancer Institute. The proceedings include not only the papers but also the major contributions to the discussion.

The treatment is primarily biochemical—all the major research groups in this field (American, Belgian, British, Czech, Dutch, French, German, Hungarian, Israeli, and Japanese) were represented—and the book provides the first authoritative and complete account of progress in this rapidly expanding field. It will be read by biochemists but also by those cytologists, biophysicists, and chemists who are interested in such topics as the function of nucleic acids, the structure of cytoplasm, the properties of ribosomes, the biosynthesis of viruses, or the synthesis of antibodies.

Biological Approaches to Cancer Chemotherapy

A Symposium held at Louvain, June 1960, under the auspices of UNESCO and the World Health Organization

Edited by R. J. C. HARRIS, *Imperial Cancer Research Fund, London*

April 1961, 431 pages, 100s.

These collected papers, with the discussion relevant to them, will be of interest to all those engaged in the study of cellular activities. It is becoming increasingly apparent that the "malignant cell" is not usually distinguishable from its normal counterpart by any *gross* biochemical or ultrastructural property. The problem of destroying such cells specifically by chemotherapy may therefore ultimately depend upon a greater knowledge of small differences between the normal and the malignant cell. Many of the possibilities towards this end are brought together here.

Thus the biochemists outlined possible ways of introducing cell specificity into cell poisons. They also discussed cellular detoxication mechanism; lysosomes in chemotherapy; feedback in enzymic pathways for nucleic acid metabolism; "lethal" synthesis; and drug design, especially in the fluoropyrimidine and in the nitrogen mustard fields.

The biologists considered variations in tumour cell populations; pinocytosis; immunological properties of tumour cells; tissue culture "malignancy"; hormones and hormone-dependence; oncogenic and oncolytic viruses; differentiation; and cell ultrastructure.

ACADEMIC PRESS, London and New York

CARCINOGENESIS

By

L. HIEGER

Chester Beatty Research Institute
Royal Cancer Hospital, London

February 1961, 138 pages, 35s.

The principal theme of this book is to put into perspective the theories which have been advanced to interpret the process whereby normal cells and tissues are transformed into cancer.
The basic concept of carcinogenesis is discussed in a critical survey of the theoretical approaches to this complex problem which involves cytological, genetic, biochemical, chemical, and physical aspects. Experimental studies on chemical carcinogens are incorporated in the well-documented analysis of the theories.

Contents: Theories of Carcinogenesis
Immunological Theory of Carcinogenesis
Carcinogens
Carcinogenesis and Chemical Structure
Carcinogenesis as a Multi-Stage Process
Work of Earle, of Warburg and of Cowdry
References—Author Index—Subject Index

In preparation

HORMONES IN BLOOD

Edited by

C. H. GRAY, *King's College Hospital Medical School, London*
and A. L. BACHARACH, *Formerly of Glaxo Laboratories Ltd.*

July 1961, about 656 pages, approx. 120s.

Methods of hormone assay in blood and plasma are presented in a series of informative and fully documented chapters by scientists expert in the specific endocrine substances discussed. The book gives a critical survey of the different hormones, the methods of their determination in blood, and supplementary information on their presence in other body fluids. This material has been integrated with a summarization of the chemical contributions and physical properties of the hormones, their sources and biosynthesis, and their secretion and fate in the organism.

ACADEMIC PRESS, London and New York

GENERAL AND COMPARATIVE ENDOCRINOLOGY

An International Journal

Editors-in-Chief: AUBREY GORBMAN, *Columbia University*
and E. J. W. BARRINGTON, *The University, Nottingham*

Volume 1, 1961 (6 issues) $18.00 128s. 6d.

GENERAL AND COMPARATIVE ENDOCRINOLOGY will be devoted to the publication of studies on cellular mechanisms of hormone action, and on functional, developmental, and evolutionary aspects of vertebrate and invertebrate endocrine systems. It will serve as a central forum for original research papers in endocrinology, but occasional brief reviews dealing with particular fields or problems are invited.

Manuscripts may be submitted to either Editor.

Subscription orders should be sent to the publishers.

Contents of Volume 1, Number 1, April 1961

Interrelationships of the Thyroid and Pituitary in Embryonic and Premetamorphic Stages of the Frog, *Rana pipiens*
By NANCY WEBER KAYE

Induction of Ovulation *in Vitro* in *Rana pipiens* with Steroids
By PAUL A. WRIGHT

The Mode of Action of the Heart Accelerating Factor from the Corpus Cardiacum of Insects
By K. G. DAVEY

Evidence for the Presence of Arginine Vasotocin (8-Arginine Oxytocin) and Oxytocin in Neurohypophyseal Extracts from Amphibians and Reptiles
By WILBUR H. SAWYER, ROBERT A. MUNSICK, *and* H. B. VAN DYKE

Demonstration of Aldosterone and Corticosterone as the Principal Steroids Formed in Incubates of Adrenals of the American Bullfrog (*Rana catesbeiana*) and Stimulation of Their Production by Mammalian Adrenocorticotropin
By HANS CARSTENSEN, ANTON C. J. BURGERS, *and* CHO HAO LI

Karyometric Studies of the Preoptic Nucleus of *Rana temporaria*
By K. DIERICKX *and* N. VAN MEIRVENNE

The Effect of Hyperthyroidism on the Uptake of I^{131}-labeled Triiodothyronine by the Brain and Pituitary of the Male Rat
By DONALD H. FORD

Hypothalamo-Pituitary Relations in Metamorphosis of *Ambystoma*
By WILLIAM ETKIN *and* WALTER SUSSMAN

Vasopressin and Oxytocin in the Mammalian Hypothalamus
By K. LEDERIS

ACADEMIC PRESS, New York and London
111 Fifth Avenue, New York 3 17 Old Queen Street, London S.W.1

Complete in 5 volumes

THE CELL

Biochemistry, Physiology, Morphology

Edited by JEAN BRACHET, *Université libre de Bruxelles*
and ALFRED E. MIRSKY, *The Rockefeller Institute*

Volume 2:

CELLS AND THEIR COMPONENT PARTS

July 1961, 916 *pp.*, $25.00 178*s.* 6*d.*

The Cell Membrane and Its Properties
By ERIC PONDER, *Nassau Hospital, Mineola, New York*

Plant Cell Walls
By K. MÜHLETHALER, *Technische Hochschule, Zürich*

Ameboid Movement
By ROBERT D. ALLEN, *Princeton University*

Cilia and Flagella
By DON FAWCETT, *Harvard Medical School*

Mitochondria (Chondriosomes)
By ALEX B. NOVIKOFF, *Albert Einstein Medical College, Yeshiva University, New York*

Lysosomes and Related Particles
By ALEX B. NOVIKOFF, *Albert Einstein Medical College, Yeshiva University, New York*

The Chloroplasts: Inheritance, Structure, and Function
By SAM GRANICK, *The Rockefeller Institute*

Golgi Apparatus and Secretion Granules
By A. J. DALTON, *National Cancer Institute, Bethesda, Maryland*

The Ground Substance; Observations from Electron Microscopy
By KEITH B. PORTER, *The Rockefeller Institute*

The Interphase Nucleus
By ALFRED E. MIRSKY, *The Rockefeller Institute, and* SYOZO OSAWA, *Nagoya University*

Nucleocytoplasmic Interactions in Unicellular Organisms
By JEAN BRACHET, *Université libre de Bruxelles*

AUTHOR INDEX—SUBJECT INDEX.

Volume 3:

MITOSIS, AND MEIOSIS

September 1961, *about* 500 *pp.*, *approx.* $15.00 107*s.* 6*d.*

Mitosis and the Physiology of Cell Division
By DANIEL MAZIA, *University of California, Berkeley*

Meiosis
By MARCUS RHOADES, *Indiana University*

AUTHOR INDEX—SUBJECT INDEX.

ACADEMIC PRESS, London and New York

Journal of Theoretical Biology

Chief Editor: J. F. Danielli

VOLUME I, 1961

Number 2, April

LINDEGREN, CARL C., The Biological Function of Deoxyribonucleic Acid

BERGNER, P.-E. E., Tracer Dynamics : 1. A Tentative Approach and Definition of Fundamental Concepts

WILLMER, E. N., Human Colour Vision and the Perception of Blue

PERAULT, A.-M., VALDEMORO, C. and PULLMAN, B., Electronic Aspects of the Mechanism of Action of Xanthine Oxidase

CLARINGBOLD, P. J. AND BARKER, J. S. F., The Estimation of Relative Fitness of Drosophila Populations

CAIANIELLO, E. R., Outline of a Theory of Thought-Processes and Thinking Machines

VANDER, ARTHUR J., WILDE, WALTER S. AND MALVIN, RICHARD L., A Theoretical Mode of Action of Aldosterone

YČAS, MARTYNAS, Replacement of Amino Acids in Proteins

CALVIN, MELVIN, Quantum Conversion in Photosynthesis

Number 3, July

SHAPIRO, HERBERT, On the Relation between Ionization and Biological Activity in some Sympathomimetic Amines

MARLER, PETER, The Logical Analysis of Animal Communication

RUMMEL, W., PFLEGER, K. AND SEIFEN, E., Experimental Examination of the Hypothesis of Coupling between the Sodium and Potassium Pump in Erythrocytes

WENNER, ADRIAN, M., Division of Labor in a Honey Bee Colony—a Markov Process?

POLLARD, ERNEST, Physical Aspects of Protein and DNA Synthesis

PLATT, J. R., Properties of Large Molecules that go beyond the Properties of their Chemical Sub-Groups

BERGNER, PER-ERIK E., Tracer Dynamics : II. The Limiting Properties of the Tracer System

SUGITA, MOTOYOSI, Functional Analysis of Chemical Systems *in vivo* using a Logical Circuit Equivalent

LEWONTIN, R. C., Evolution and the Theory of Games

ARNOLD, WILLIAM, Wave Mechanics and Muscular Contraction

Letter to the Editor :—

The Effect of Concentration and Rate of Intravenous Injection of Thiopentane, L. NAFTALIN AND A. STEPHENS

The price of this volume is 121*s.* 6*d.* ($17.00) *or* 86*s.* ($12.00)
for private subscription

ACADEMIC PRESS, LONDON AND NEW YORK

CONTENTS

MILLER, I. R., The Structure of DNA and RNA in the Water-Mercury Interface 229

LAMFROM, H., Factors Determining the Specificity of Hemoglobin Synthesized in a Cell-Free System 241

ROBICHON-SZULMAJSTER, H. DE, Sur un nouveau Mécanisme d'Oxydoréduction, appliqué à la Réaction d'Epimérisation: Uridinediphosphogalactose ⇌ Uridinediphosphoglucose 253

HUEHNS, E. R. and SHOOTER, E. M., The Polypeptide Chains of Haemoglobin-A_2 and Haemoglobin-G_2 257

BEER, M., Electron Microscopy of Unbroken DNA Molecules . . . 263

KOZINSKI, A. W. and UCHIDA, H., Phage DNA Subunits in the Phage Precursor Pool. 267

THOMAS, C. A., JR. and BERNS, K. I., The Physical Characterization of DNA Molecules Released from T2 and T4 Bacteriophage. 277

BERNS, K. I. and THOMAS, C. A., JR., A Study of Single Polynucleotide Chains Derived from T2 and T4 Bacteriophage 289

BRADBURY, E. M., PRICE, W. C. and WILKINSON, G. R., Infrared Studies of Molecular Configurations of DNA 301

Review Article—

JACOB, F. and MONOD, J., Genetic Regulatory Mechanisms in the Synthesis of Proteins 318

Letters to the Editor—

Temperature Dependence of the Adsorption of Native DNA in a Polarized Water-Mercury Interface, I. R. MILLER 357

Comparative Studies on Tryptic Digests of Myoglobins, A. STOCKEDL . 362

Printed in Great Britain at the Aberdeen University Press